Understanding the diversity of maternal microbiota species *Bifidobacterium* using culturing and genomic approaches

Rhianna-lily Smith

A thesis submitted for the degree of Masters by Research (MSc by Research) in Biological sciences

2023

University of East Anglia
School of Biological Sciences

Quadram Institute Bioscience
Gut Microbes and Health

Supervisor
Professor Lindsay Hall

Title: Understanding the diversity of maternal microbiota species *Bifidobacterium* using culturing and genomic approaches

Research Hypothesis:

There are a variety of *Bifidobacterium* species and strains which are present in the maternal microbiota that encode a range of beneficial functional traits.

Abstract:

The establishment of the gut microbiota starts at birth, and colonisation of this early-life microbial community has a significant impact during infancy and in later life. It is now appreciated that the maternal microbiota also plays a vital role during pregnancy, with rapid changes in the microbiota composition observed throughout the gestation period. Disturbances in the maternal microbiota have also been identified as a risk factor for many diseases and complications during pregnancy. There has been limited work on the maternal microbiota and its composition, particularly an in-depth characterisation of beneficial genera such as *Bifidobacterium*. Recently, studies have suggested this bacterial group may modulate maternal health and foetal development, and therefore this project aims to identify which species and strains of *Bifidobacterium* are present in the maternal microbiota throughout healthy pregnancy, using a combination of culturing techniques and whole genome sequencing. A total of 140 *Bifidobacterium* colonies were cultured and isolated from 12 pregnant women during 3 timepoints during gestation and 1 post birth. Five *Bifidobacterium* species were found, with the majority identified in *Bifidobacterium longum* and *Bifidobacterium animalis.* The genomes were run through several bioinformatics programmes to identify and visualise both the phylogenetic relationships between the isolates, and several known human milk oligosaccharide clusters and antimicrobial resistance genes; which were found in all 5 species. Although many of the *Bifidobacterium animalis* strains were shared among participants, all *Bifidobacterium longum* strains were participant specific. This research offers essential insights into *Bifidobacterium* species and strains that reside in the maternal gut, and their overall diversity and functional traits that they encode. This data presented here may allow additional studies to further our understanding on how this genus contributes to maternal wellbeing.

Table of Contents:

List of Figures

List of Tables

Supplementary Tables

Abbreviations

HMOs - Human Milk Oligosaccharides

WGS - Whole Genome Sequencing

HGT - Horizontal Gene Transfer

SCFAs – Short Chain Fatty Acids

C-section - Caesarean section

GOS – Galacto-oligosaccharides

FDA – Food and Drug Association

LBP – Live Biotherapeutic Products

FMT – Faecal Microbiota Transplant

CDI – *Clostridioides Difficile* Infection

IBD – Inflammatory Bowel Disease

PE – Preeclampsia

TMAO – Trimethylamine-N-Oxide

NDOs – Non-digestible oligosaccharides

FOS – Fructo-oligosaccharides

RS – Resistant Starch

NEC – Necrotising enterocolitis

ANI – Average Nucleotide Identity

PEARL – Pregnancy and Early Life study

PCR – Polymerase Chain Reaction

BR – Broad range

RED – Relative Evolutionary Divergence

SNP – Single Nucleotide polymorphisms

ARGs - Antimicrobial Resistance Genes

qPCR – Real-time Polymerase Chain Reaction

MAGs – Metagenome-assembled genomes

LNT – Lacto-N-tetraose

LNnt – Lacto-N-neotetraose

1.0 Current Literature

1.1.1 Overview of the gut microbiota

The gut microbiota is the term that describes the 10^{13} to 10^{14} different micro-organisms that reside in the human gastrointestinal tract, (1), consisting of bacteria, fungi, archaea, and viruses. The collective microbiome (i.e. microbiota associated genes) has been estimated to contain approximately 100 times the number of genes than the human genome, (2). This complex microenvironment plays a critical role in host health, throughout an individual's whole lifespan. Notably, gut microbiota disturbances have been associated with various complications such as diabetes, inflammatory diseases, and an increased susceptibility to infections and certain cancers (3-5). Neurological conditions have also shown to be linked with disturbances of the normal gut microbiota composition via the gut-brain axis including Alzheimer's and even anxiety/depression (6, 7). Establishment of gut microbiota starts at birth with waves of microbial colonisation occurring throughout the first two years of life, before an 'adult-like' gut microbiota is established by 3-5 years of age (8). Many factors, which affect healthy gut microbial composition have been identified, including diet, stress, and the use of antibiotics, (9-11), thus our understanding of how we can use this information to aid human health is slowly developing. Given the focus of this MRes was understanding the gut microbiota during pregnancy and early life, this literature review aims to summarise the gut microbiota's role during these key life stages with an emphasis on the importance of the beneficial microbiota genus *Bifidobacterium*.

1.1.2 Main players of the Microbiota

The Human Microbiome Project, launched in 2007-2016, provided the scientific community with the first large overview of the gut microbiota and its potential roles in host health and disease. Further work has determined that between 93-99% of intestinal bacteria belong to one of four phyla's: Firmicutes, Bacteroidetes, Proteobacteria or Actinobacteria, (12), which have recently been renamed to Bacillota, Bacteroidota, Pseudomonadota and Actinomycetota respectively. Throughout a human's life their gut microbiota composition can change drastically. For example, within the infant gut there is a higher proportion of bacteria involved in the digestion of breast milk sugars – i.e. human milk oligosaccharides (HMOs) such as *Bifidobacterium* (belonging to the Actinomycetota phyla), compared with the more diverse adult-like microbiota that is structured to metabolise plant-derived polysaccharides

(represented by bacteria belonging to the Bacillota & Bacteroidota phyla's) (13); this changing microbial profile is represented in **Figure 1.1**.

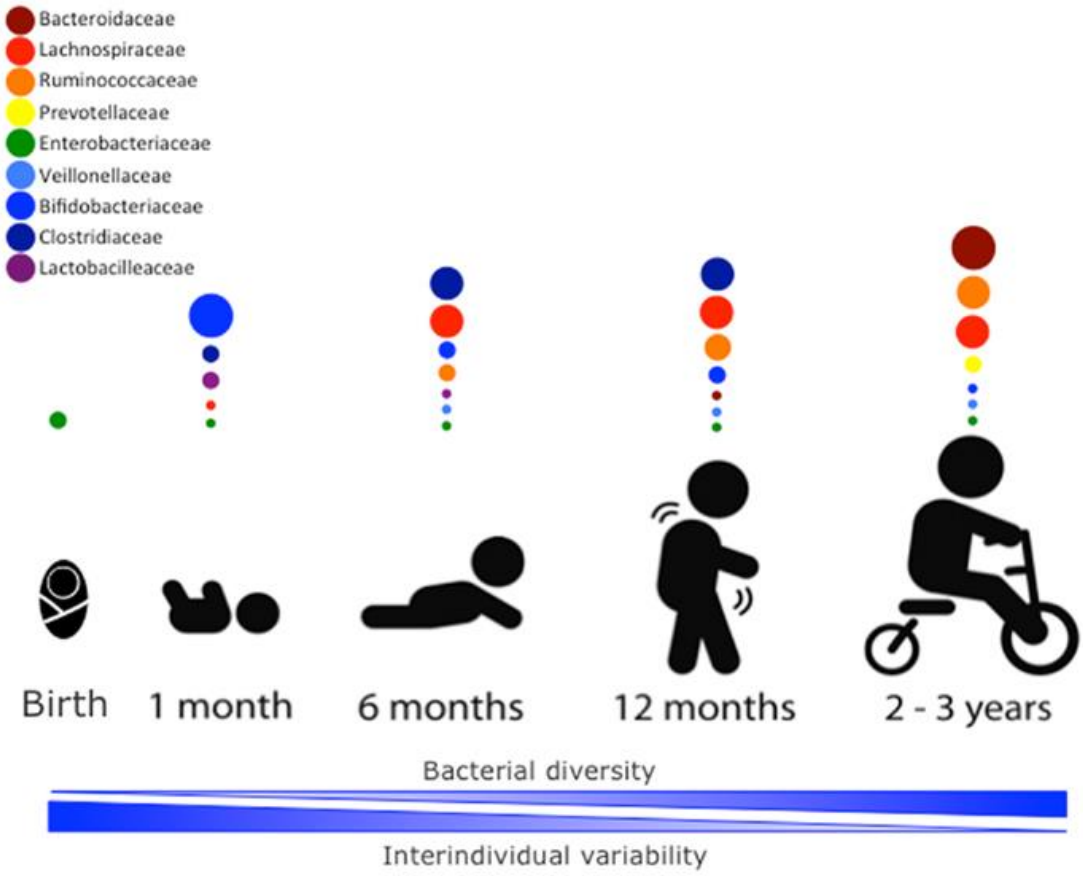

Figure 1.1: The establishment of the gut microbiota begins during birth and carries on till old age. At the infant stages a higher inter-variability is seen with a lower diversity which gradually increases until around the age of 3 years old where an adult-like microbiota profile is seen which does not change significantly through adulthood. Figure taken from(14).

1.1.3 How is the microbiota is profiled?

Currently, researchers use a variety of tools to profile the gut microbiota which include using 16S rRNA gene amplicon sequencing to identify bacterial genera, and more recently the growing field of metagenomics also provides insights down to the species and strain level and functional potential. Although up until recently it was considered an older technique, culture or 'culturomics' has become more widely used once again, coupled with whole genome sequencing (WGS), and as isolates are obtained, this also allows potential next stage therapy development.

Studying bacterial diversity and evolution, including gut microbiota members, is key for improving our understanding of their importance in human health and disease. In general, bacterial diversity concerns species richness, and functional and phylogenetic diversity, which within the gut microbiome links to ecosystem stability and effects on host health, (15). The human gut microbiota has been shaped throughout generations by evolution and adaptions to the environment. These changes in gut composition have been studied in several populations and can be used as biomarkers for gut health.

One aspect to consider when studying the specific microbiota members are the differences which occur in the pangenome compared to the core genome within a species. The pangenome concerns the collective genes which exist in a given species of interest, (16), developed by Tettelin et al in 2005, (17), and can be used to represent genome evolution through gene loss and gain. The pangenome can be further broken down into the core genome, those which are shared by all the species in a clade, and the accessory genes, those which may only be found in some species/isolates, (18). The core genome offers insight into genes and clusters which have remained in species for evolutionary benefit and can also be used to understand how a particular species contributes to its host. Horizontal gene transfer (HGT) is the transfer of genetic material between organisms which doesn't occur through a parent to offspring relationship, (19), and also plays a vital role in the evolution of many organisms. HGT can occur through many mechanisms, including transformation, transduction, and conjugation, (20), and this must be taken into account when studying the phylogenetic relationship between bacteria.

1.1.4 The Microbiota and human health

The composition of the gut microbiota can contribute positively to human health, which can generally be characterised into three 'functions': nutrient processing/production, immune development, and infection defence. However, disturbances in healthy microbiota signatures have been associated with a plethora of diseases and neurological conditions as stated above. More specifically, the microbiota is heavily involved in the digestion and break down of complex carbohydrates, proteins, and some fats. Pyruvate is produced during the breakdown of carbohydrates, which is then further broken down into succinate, lactate, or acetyl-CoA, to generate energy for their host (21). Pyruvate intermediates are further metabolised to produce microbial products that can affect their host. It's been widely documented that certain microbes

residing in the gastrointestinal tract can synthesise nutrients for their host such as vitamin K, and B group vitamins, (22), which play a role in certain physiological pathways. Bacteria in the gut also produce short chain fatty acids (SCFA's) such as acetate, propionate, and butyrate, (23). The production of these bacterial metabolites has been implicated in several areas of host wellbeing, including reduced levels observed in inflammatory bowel disease, (24), and even the cause of a change of mood and behaviour due to their ability to influence the integrity of the blood–brain barrier through signalling of the gut-brain axis, (25); their modes of action are summarised in **Figure 1.2**.

The gut microbiota plays a critical role in the development and modulation of the immune system. Germ-free mice (i.e. those that lack all microbes) exhibit impaired immune development, however colonisation of intestinal bacteria can restore these immune system disturbances, (26). The way the gut microbiota exhibits its effect on the immune system occurs through many pathways including the modulation of T cells and the development and function of macrophages, (27, 28). It is also widely accepted factors which affect the gut microbiota composition during the early-life stages have a knock-on effect on the immune system during adulthood; for example babies delivered via a caesarean-section (C-section) have an increased risk of immune disorders such as asthma, and Type 1 diabetes, (29, 30). Furthermore, the production of bacterial metabolites (such as SCFAs), as mentioned above, also contribute to the modulation of the immune system via stimulation of specific cellular pathways.

One way the gut microbiota contributes to infection defence (i.e. colonisation resistance) is through the competition for nutrients against incoming microbes, and preventing colonisation of pathogens, (31). Certain commensal bacteria also secrete (antimicrobial) inhibitory substances, such as *Streptococcus salivarius,* which secretes bacteriocins that targets *S. pneumoniae* , (32).

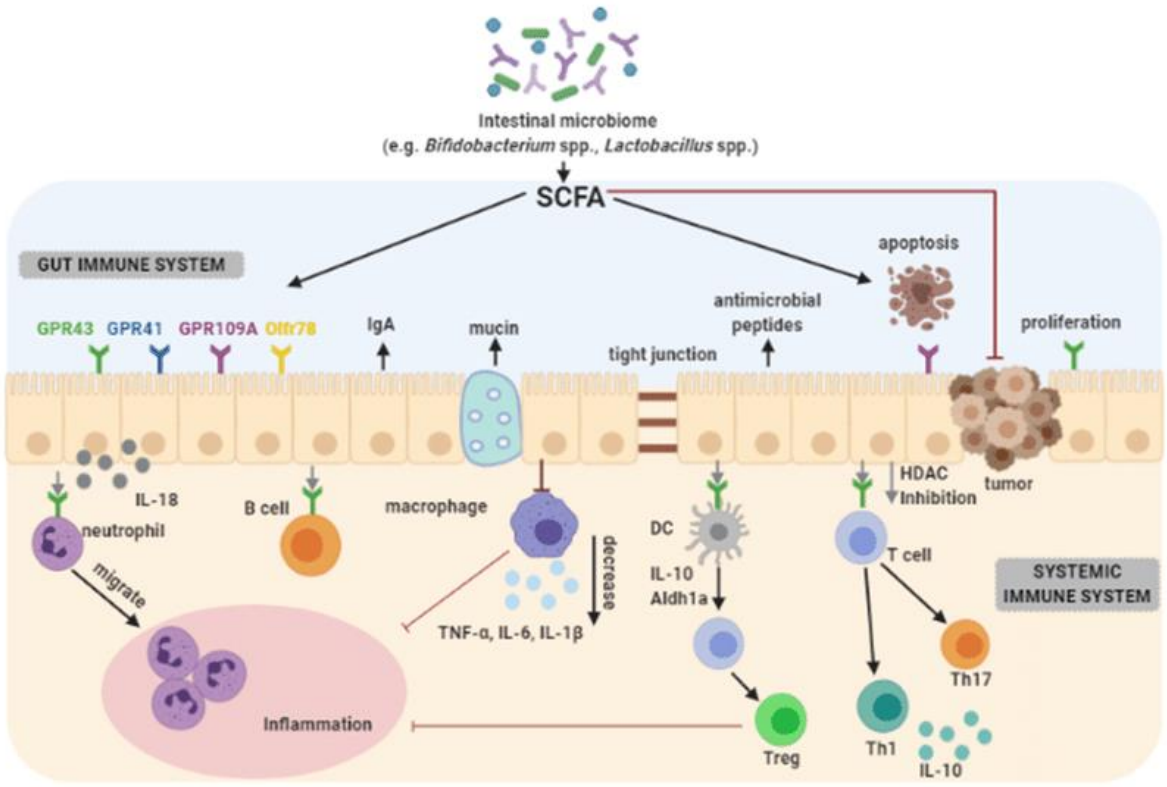

Figure 1.2: SCFAs are a bacterial metabolite that play a role in many different pathways to aid the immune system. SCFAs modulate the intestinal barrier function by inducing intestinal epithelial cell secretion of interleukins, mucin, antimicrobial peptides, and upregulating the expression of tight junctions. They also regulate the T cell function through G-protein-coupled receptors. Figure taken from (33).

1.1.5 *Clinical uses of the Microbiota*

As our understanding of the benefits of a healthy microbiota composition increase, as does the development of therapeutic strategies which work to target this system. One technique used to modify the gut is through the use of pre and probiotics. Prebiotics are defined as 'a selectively fermented ingredient that results in specific changes in the composition and/or activity of the gastrointestinal microbiota, thus conferring benefit(s) upon host health' (34). These include dietary components such as galactooligosaccharides (GOS) and breast milk derived HMOs, which can stimulate the growth of *Bifidobacterium*, *Lactobacillus*, and *Bacteroides*, (35). Probiotics are defined as 'live microorganisms which when administered in adequate amounts confer a health benefit on the host', (36). The use of probiotics for promoting a healthy gut microbiota has been heavily debated, however their effectiveness in treating conditions such as acute diarrhea in children has been well documented, (37). Daily intake of milk-based dietary probiotics has also been shown to reduce blood pressure during pregnancy in

Primiparous Women, also known as first time mothers, (38). However, the benefits of probiotics are strain specific, and therefore more research is necessary to understand which bacterial species and strains are more effective (coupled with the 'right' diet).

In 2010 the Food and Drug Administration (FDA) created a new guideline for products that contained live microorganisms which provided a benefit to their host through the prevention, treatment and or cure for a disease or condition, and by 2012 this category of medicinal products was termed Live Biotherapeutic Products (LBPs) (39). The new guideline has allowed for more stringent regulation on these types of products and further research into their therapeutic applications.

Fecal microbiota transplants (FMT) involve the transfer of healthy intestinal microbiota, from a donor to a diseased individual to aid the restoration of their microbiota composition. Although FMT in a clinical setting is still not widely used for many diseases, many trials have shown great promise for its potential. However, for the treatment of recurrent *Clostridioides difficile* infection (CDI), FMT has an efficacy rate of >90%, (40), and it is now recommended on the UK NICE clinical guidelines; there has also been evidence of its ability in the treatment of inflammatory bowel disease (IBD) (41). However even with this knowledge, FMT still remains an unstandardised treatment and its risks need to be carefully considered before it is more widely accepted.

1.2.1 The general principles of pregnancy

Although the early-life microbiome has been well characterised, research on the maternal gut microbiota is still limited. In this project, I have explored the bacterial species (with a focus on *Bifidobacterium*) which are present during the gestation period, thus the following information provides details on this period of life. During pregnancy the body goes through a substantial number of physiological, metabolic, and hormonal changes, affecting all organs of the body, (42). From an increase in adipose tissue to a reduction in insulin sensitivity, the body prepares and nurtures the growth of the foetus, (43). These physical changes result in a significant change in maternal metabolism which is vital for a healthy pregnancy, (44). A large portion of these switches are driven by the many endocrine pathways associated with pregnancy; **Figure 1.3** outlines the main hormones involved in this process. The composition of the gut microbiota

also differs in pregnant women, (45), with an increase in bacterial abundance observed over the gestational period, (46).

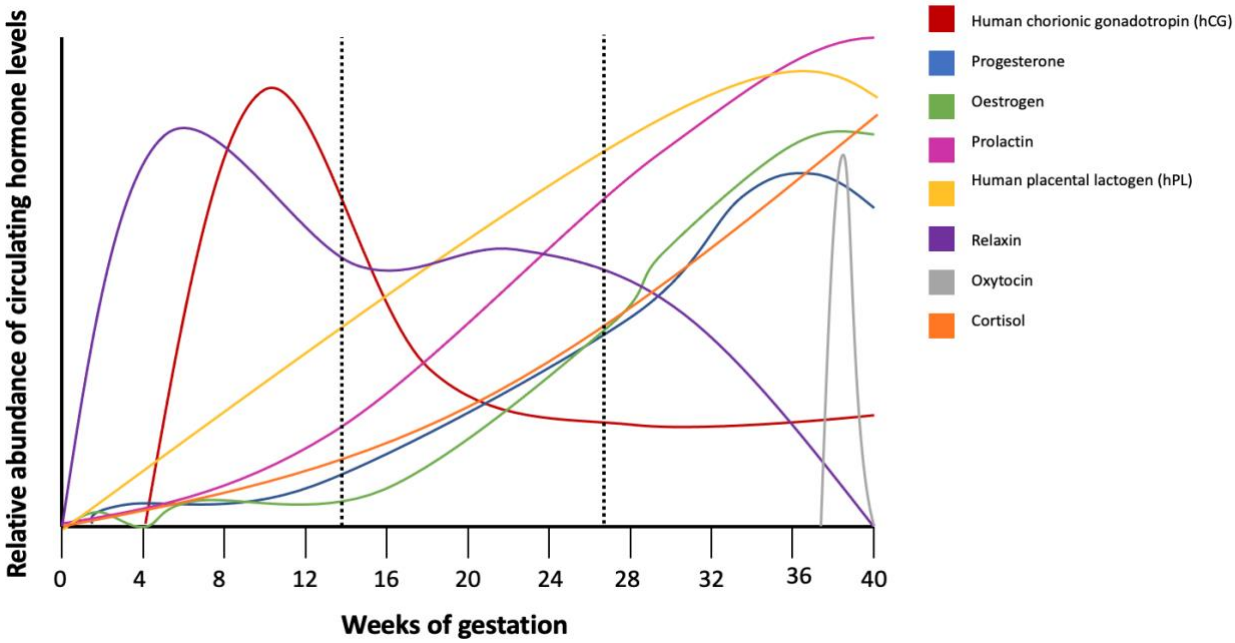

Figure 1.3: Relative abundance circulating hormone levels of 8 different hormones which drive various changes throughout pregnancy during the 40 weeks of gestation split into the three trimesters.

1.2.2 The Maternal Microbiota

It has only recently been acknowledged that, as well as the many physiological changes observed during pregnancy, there is also a dramatic change in the composition of gastrointestinal bacteria. During the first trimester, the gut microbiota profile typically represents that of a non-pregnant woman, (47), however as the gestation period progresses, more and more changes can be observed that are not solely associated with health status or diet, (44). There is an overall increase in the relative abundance of Pseudomonadota and Actinomycetota, including *Bifidobacterium*, and a reduced abundance of Bacillota and Bacteroidota, (44). One explanation for the change in microbiota composition at various stages of pregnancy could be due to the constant hormonal fluctuations that occur. Both estrogen and progesterone can affect the gut microbiota composition through their activity on bacterial metabolism and growth, (48). At trimester 3 the abundances of health-related bacteria is also affected, (44). One mouse study discovered that when bacteria, taken from the trimester 3 microbiota, was transferred to germ-free mice there was a greater increase in adiposity and insulin insensitivity compared to when trimester 1 microbiota was used, (44). These changes

in composition have a direct effect on the levels of metabolites circulating. SCFA levels studied in the cecum in rats found that the abundance of acetic and propionic acid was increased in pregnancy, as were the levels of butyric and caproic acid, (47). These distinct changes could have implications on pregnancy outcomes and/or infant complications.

1.2.3 Factors which affect the Maternal Microbiota

Factors that can impact the early-life microbiota have been well documented, such as mode of delivery, type of feeding and length of gestation, (49-51). However, as more research reveals how the maternal gut microbiota can affect the result of pregnancy, there is a growing demand to understand what factors may also alter the maternal microbiota during the gestation period. One factor that has been identified to cause gut perturbations in pregnant women is obesity, (52). Studies have found that levels of *Bacteroides* and *Staphylococcus*, in the faeces, are increased in overweight pregnant woman compared to those of a healthy weight, (46). Levels of Enterobacteriaceae, such as *Escherichia coli* have also been reported to be elevated in overweight pregnant women compared to their healthy weight counterparts, (53). Mice experiments suggest endotoxemia and inflammation, induced by obesity, may cause the increase in Gram-negative bacteria seen in these women, (54). Another factor that has been observed to alter the microbiota in pregnant women is blood pressure. The abundance of *Odoribacter*, a butyric acid–producing bacterium, is negatively correlated with systolic blood pressure in during early pregnancy, (55). However, an increased systemic blood pressure is also associated with an increased weight gain, therefore providing more evidence for a pregnant woman's BMI being a driver of disturbances in the maternal gut. Antibiotics administered during pregnancy can also cause large perturbations in the maternal gut, which is covered in more detail below.

Although these studies have provided some insights into pregnancy specific perturbations, further research is needed to understand how a mother's choices and/or experiences during pregnancy could impact her gut microbiota composition and therefore affect the growing foetus. However, it must be noted that these studies are only correlation studies and have not shown causation.

1.2.4 The Maternal Microbiota and disease

The composition of the maternal microbiota and associated detrimental changes have been associated with several diseases and complications during pregnancy. Preeclampsia (PE) is a pregnancy complication characterised by hypertension and evidence of organ damage after the 20th week of gestation, (56). Although the pathophysiology of PE is not yet fully understood, Lv et al. discovered a link between the maternal microbiota and the condition, (57). They found the composition of the gut microbiota greatly differed from healthy pregnant women to that of women suffering from early onset PE for example, 8 bacterial genera, including *Blautia, Ruminococcus, Bilophila,* were significantly increased in patients with early onset PE compared to the healthy controls. However, 5 genera were also depleted in the PE group including *Faecalibacterium.* A second study suggested the gut microbiome may exaggerate the inflammatory response by affecting the production of proinflammatory metabolites resulting in PE such as Trimethylamine-N-Oxide (TMAO) through the downregulation of IL-10, (58). Hu et al, (59), showed the low serum levels of the SCFA acetate was also associated with PE however a diet high in fibre to promote SCFAs production, did result in a decrease in the incidence of PE.

Disturbances of the mother's microbiota has also been associated with disruption with the infant's gut microbiome and complications later in life, (60). For example mothers, with a high increased weight and BMI during pregnancy, have shown to give birth to infants with a fecal microbial composition that is different compared to those born from mothers with a healthy BMI, (61). They observed a lower abundance of benefcial gut bacteria, such as *Bifidobacterium*, and higher levels of *Staphylococcus,* (62), which is linked to inflammatory diseases, (63).

The use of antibiotics during pregnancy has also been reported to cause a susceptibility to many later life conditions. Almost 80% of all medications administered to pregnant women during the gestation period are antibiotics, (64), and their disruption to the human gut microbiota is well established, (65), and now their effect on the infant microbiota is emerging. Maternal use of antibiotics has been linked to a 1.3- fold increased risk of asthma in their offspring, (66), and even associated with neurological disorders, (67).

1.2.5 Vertical transmission between the Maternal and Early life Microbiota

Vertical transmission describes the biological process of transfer of microbes from the maternal microbiota to the infant. *In utero* is a sterile environment, and vertical transmission doesn't take place until during birth as the baby passes through the birth canal and also comes into contact with the maternal gut microbiota, (68). How these maternal microbes influence foetal health, prior to birth is potentially through bacterial metabolites, as certain bacterial products are able to cross the placenta and effect the growing foetus and may even affect the neural development (69). For vertical transmission – which includes passage of maternal *Bacteroides* and *Bifidobacterium* to infants, appears to play a significant role in early life microbiota structuring and overall infant development, (70). When the normal process of vertical transmission is disrupted, such as when a C-section is preformed, alterations are observed in the infant microbiome, (30). Infants delivered via C-section are colonised by skin bacteria and microbes present in the hospital environment, (71), compared with those born naturally which are predominantly colonised by bacteria from the maternal vaginal and gut microbiota, (72). These disturbances that occurs have been associated with a difference in adaptive and innate immunity in these infants, with lower levels of IgM, IgA, and IgG secreting cells observed in individuals delivered via C-section even up to the age of 2 years old, (73). Several clinical trials have set out to try and restore the normal microbiota composition through Maternal Fecal and vaginal Microbiota Transplantation, (74, 75), many of which found restorative ability in this technique, however there is much more research needed in this area to understand the long term effect of microbiota transfer.

1.3.1 Overview of Bifidobacterium

Bifidobacterium are considered a beneficial or 'probiotic' bacterial genus, belonging to the Actinomycetota phylum (76). Currently to date, over 100 species have been identified, which can be split into four phylogenetic groups: *Bifidobacterium longum*, *Bifidobacterium pseudolongum*, *Bifidobacterium animalis*, and *Bifidobacterium thermacidophilum*, (77). *Bifidobacterium* was first isolated from the faeces of breast-fed infants in 1899, (76). *Bifidobacterium* are Gram-positive, mostly anaerobic bacteria with a genome size varying between 1.73 to 3.25 Mb and a high G+C content ranging of 55% to 67%, (78). The genus is recognised for the Y-shape morphology, also known as bifid, and lack of flagella, (78). Typically found in the gastrointestinal tract, *Bifidobacterium* have also been reported to be present in human blood, sewage, and food products, although it's worth noting these could

have been due to contamination from the gut. These microbes have been linked with many vital biological processes in the human body, making them valuable for human life.

1.3.2 Bifidobacterium and human health

One of the ways that *Bifidobacterium* contributes to human health is via carbohydrate metabolism. Non-digestible oligosaccharides (NDOs) are a group of glycans which include various prebiotics, such as fructo-oligosaccharides (FOS) and GOS, (79), with products produced during carbohydrate metabolism aiding human health.

Between 12-14% of the *Bifidobacterium* genome encodes for carbohydrate metabolism, (80), allowing them to make a large metabolic contribution towards their host. The presence of natural carbohydrates, such as fibres like resistant starch (RS) found in fruit, beans, and bread, stimulate the growth of *Bifidobacterium*, (81). Through fermentation, these *Bifidobacterium* strains break down non-digestible carbohydrates to produce SCFA's, discussed below. High levels of *Bifidobacterium* may also correlate positively with the an improved glucose-induced insulin secretion and glucose tolerance, (54), through a reduction of inflammation, which could have huge applications for the pre-diabetic population.

Another very important role of *Bifidobacterium* is their ability to produce SCFAs and lactate. SCFAs can regulate the immune system through dendritic and T-cells as well as their role in cytokine production inhibition, (82). SCFAs also reduce the pH of the gastrointestinal tract, allow for a higher availability of calcium and magnesium, and can aid in the inhibition of pathogenic bacteria such as *Helicobacter pylori* (83). As well as SCFAs, some strains can also produce vitamins such as biotin, riboflavin, thiamine, folic acid, and nicotinic acid, (84-86). Many of these vitamins can only be produced by bacteria, fungi, and plants, highlighting the importance of having *Bifidobacterium* present in the gut for human health.

Certain *Bifidobacterium* species have been reported to produce bacteriocins, antimicrobial peptides which can kill or inhibit other bacteria without harming the original bacteria, (87). For example, the bacteriocin Bifidin has shown activity against both *E. coli* and *Staphylococcus aureus,* (88) and a more recent study identified a novel bacteriocin called Bifidococcin_664 which reported immune stimulatory effects and activity against the pathogen, *Clostridium perfringens* (89). There are several other ways that *Bifidobacterium* provide a protective ability to their hosts, such as production of acid and the degradation of oxalate (90). *Bifidobacterium* has also been shown to be important for modulation of the gut barrier. One recent study found

mice which were administered *B. breve* had ~4000 genes upregulated, many of which have been identified to playing a role in epithelial barrier function, compared to those which were administered PBS, (91). All these factors added together show the importance of *Bifidobacterium* in the gut.

1.3.3 Bifidobacterium and Maternal and Early life gut health

Bifidobacterium are the first colonisers in the infant gut microbiota and dominate the gut for the first year of life, (92). These microbes has been shown to offer protection against diseases for the infant such as celiac disease, asthma and obesity, (93, 94), and for the mother during pregnancy, (95). Although there is evidence of vertical transfer of these species between mother and infant, the most frequently seen in both include *B. bifidum* and *B. breve*, which have been observed in the infant gut up to the age of one, (96). Distinct infant and adult *Bifidobacterium* species are still to be classified, although the bifidogenic effect has been reported to be strain specific, promoting the growth of *B. longum* subsp. *infantis* over others, (96), suggesting certain strains may be more favoured in the infant to adult microbiota and provides support for the idea that particular *Bifidobacterium* species occupy different niches in their host. One recent study used germ-free and specific-pathogen free species mice to show *B. breve*, affects the maternal body composition and feto-placental growth; it was suggested this could be through altered metabolites in the mother, (97). Many studies have shown in importance of gut microbiota dominated by *Bifidobacterium* during the early-life stages and why their presence in the maternal gut is equally important.

Further research into which strains may provide specific benefits to their hosts is needed to further promote the health and mothers and their infants. *Bifidobacterium longum* and *Bifidobacterium animalis* are both commonly used as probiotic strains, (98), and can be found in the human gut microbiome. *B. longum* is split into three subspecies: *longum, infantis* and *suis*, (99), with *B. longum* subspecies *infantis* being extensively studied in the early life gut. This subspecies has been identified in aiding the maturation of the immune response and improving intestinal barrier function, (100). *B. longum* subsp. *infantis* also contains many of the enzymes required to break down HMOs, (101), discussed in more detail below. As mentioned above, *B. breve* has also been identified as important in both the maternal and infant gut.

1.3.4 Bifidobacterium and Human Milk Oligosaccharides

Human milk contains many soluble oligosaccharides called HMOs. These are made up of five basic monosaccharides: glucose, galactose, N-ethylglucosamine, fucose and sialic acid, shown in **Figure 1.4**, which follow a basic structure. Over 200 distinct HMO structures have been identified which vary depending on the individual, (102). These sugars reveal high levels of diversity in carbohydrate structure although the purpose of this diversity remains relatively unknown, (103). Although the role of human milk is to be the sole nutrient source for infants, many of the carbohydrates are indigestible and many of the nutritional benefits come from their ability to affect the gut microbiota. HMOs are a well-established prebiotic which promote the growth of *Bifidobacterium* described as the "bifidogenic" effect, (96). Certain *Bifidobacterium* strains have been shown to utilise HMOs through consumption via particular sets of genes, found in their genomes, known as HMO clusters, which encode for regulatory elements, ABC transporters, carbohydrate binding proteins and glycoside hydrolases, (104). The presence of these clusters has been shown to be species specific, (79), for example many studies have revealed *B. longum* subsp. *infantis* as having the most evolved capability in breaking down these structures compared to other *Bifidobacterium* species, which may offer an explanation as to why certain species are more likely to be present in the maternal/early life gut compared to others.

HMOs also offer protection to infants in ways which do not just involve the benefits from the increase in *Bifidobacterium* species and strains. HMOs have been shown to prevent pathogen adhesion due to their resemblance of glycan structures, (105). Pathogens which would usually attach themselves to epithelial cell structures may bind to HMOs resulting in them being excreted without effecting the host, due to these effects HMOs have also shown effective in the treatment of neonatal diarrhoea, (106).

HMOs may also offer benefits to the mother during lactation and even pre-natal with recent studies showing they are also circulating during pregnancy, (107). Although more research is needed on this area, there is the potential these carbohydrates could be offering a systemic effect similar to what is observed in breast-fed infants.

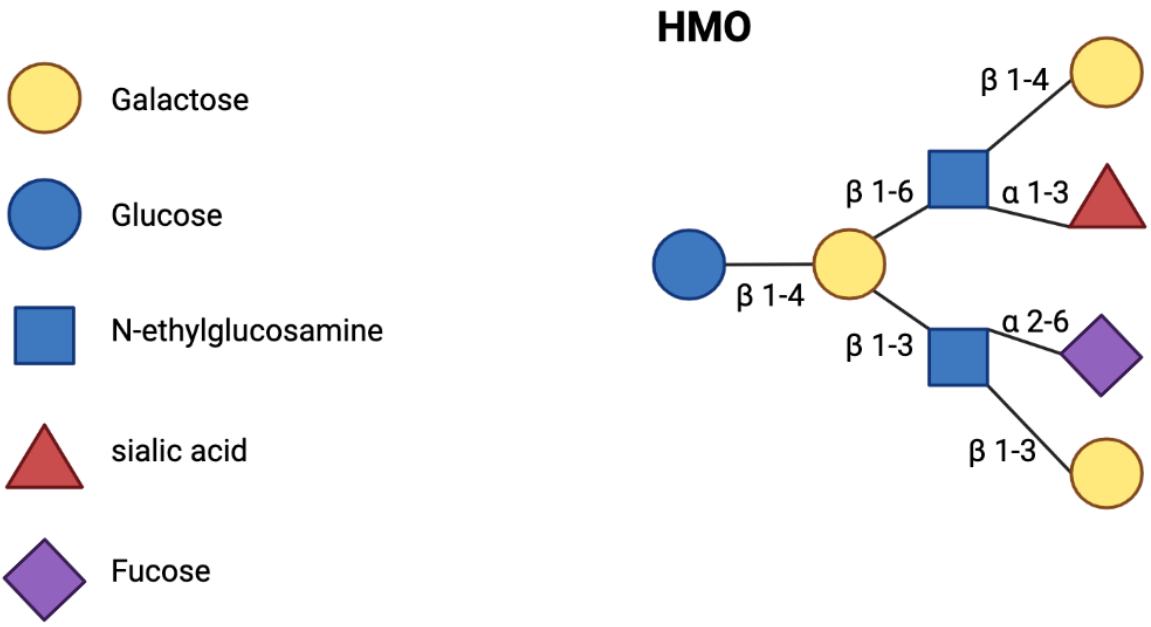

Figure 1.4: The representative structure of a HMO with the α and β linkages highlighted. These structures are lengthened by galactose and N-ethylglucosamine units which can be branched out by the further addition of sialic acid and fucose, figure adapted from, (108).

1.3.5 Bifidobacterium and pregnancy hormones

Several studies have identified that there is a dramatic change in microbial composition during the gestation period which includes and increase abundance of *Bifidobacterium,* (109), however the mechanism behind this change still remains unclear. It has been suggested that the endocrine system may be an effector of this change although there is limited research on the effect of pregnancy hormones on the growth of *Bifidobacterium*. Progesterone gradually increases throughout the 40 weeks of pregnancy, **Figure 1.3**, and one study found a direct link between increased progesterone levels and an increased abundance of *Bifidobacterium in vivo* and *in vitro*, (109). A few other studies have also suggested the impact of hormones on the gut microbiota during pregnancy, (44), however there is still a lack of conclusive evidence.

1.3.6 Clinical uses of Bifidobacterium and future use

With the range of health benefits of *Bifidobacterium* already discussed previously, many people in the scientific community have set out to utilise these in a clinical setting. One method typically used are the oral delivery of probiotics to alleviate or prevent symptoms from several diseases. Certain clinical trials have reported *Bifidobacterium* can be used to reduce the risk of antibiotic-associated diarrhoea, (110). In the context of the maternal and early-life gut

microbiota, when probiotics were administered to new-borns, who had their gut microbiota colonisation disrupted due to delivery via C-section, their gut microbiota composition profile was similar to that seen in infants delivered naturally, especially that of *Bifidobacterium* colonisation, (111). A longitudinal study looked at the effect of probiotic administration in pre-term infants with reduced *Bifidobacterium* abundance; they also reported beneficial effects in restoring the preterm infant gut microbiome and a significant reduction of the incidence of necrotising enterocolitis (NEC) (112). Many studies looking at similar objectives have also reported similar results, indicating the benefits of *Bifidobacterium* in a clinical setting. As more clinical trials are undertaken, additional uses and the potential of *Bifidobacterium* may be realised, however further studies exploring the underlying genomic traits of these beneficial taxa is key.

2.0 My project

2.1.1 Aims and Objectives

The overall objective of this project included the isolation of *Bifidobacterium* strains to allow for phylogenetic investigation and profiling of specific functional traits, more specifically with the following aims.

1. To culture, store, and isolate unique *Bifidobacterium* strains within a regulatory framework.
    a) Using fecal samples, obtained as part of the ethically approved and Quadram-BBSRC funded maternal and infant PEARL cohort, we will isolate and culture *Bifidobacterium* strains using selective agar.

2. To conduct a phylogenetic and genomic functional analysis of *Bifidobacterium* strains.
    a) The isolated strains will undergo 16S rRNA sequencing before being sent off for WGS using short- (Illumina, 60X coverage) platforms to provide high quality reference-based genomes.
    b) We will use core genome phylogeny and Average Nucleotide Identity (ANI) values to determine genetic relatedness of strains, comparing to type strains, publically available WGS, and in-house isolates.
    c) We will identify known human milk oligosaccahride (HMO) clusters, using BLAST, by comparing known protein sequences to the genomes.
    d) We will perform further functional trait analyses on the isolates, such as those concerning the presence of their antibiotic resistance profiles.

3.0 Methodology

3.1.1 Faecal Culturing and isolation

Maternal faecal samples had previously been collected at different time points throughout pregnancy as part of the ethically approved and Quadram-BBSRC funded The Pregnancy and Early Life study (PEARL) cohort. This study is concerned with understanding the importance of beneficial microbes in the development of the infant microbiota; on which supervisor (L Hall) is Chief Investigator and Sarah Phillips is study coordinator. Samples from each trimester and 1 post birth samples were used. Further details about the PEARL study can be found in the published protocol (113). The PEARL study has been reviewed and agreed by the Human Research Governance Committee at QIB and the London-Dulwich Research Ethics Committee (reference 18/LO/1703) and received written ethical approval by the Human Research Authority. IRAS project ID number 241880.

Approximately 100 mg of (frozen and then thawed) raw stool was homogenised with 1ml reduced sterile PBS. The faecal slurry was serially diluted up to 10^3 and 200 μl of the dilutions 10^2 and 10^3 were plated onto selective agar plates, containing both cysteine and mupirocin at a concentration 10^{-3} (1 ml of cysteine and mupirocin in 1000ml of BHI media), and spread using a disposable spreader until dry. Plates were left to culture in an anaerobic chamber at 37°c for up to two days.

From the dilutions, the most appropriate plate was chosen for each individual sample and five colonies, that morphologically represented *Bifidobacterium*, were picked off via a disposable inoculating loop and restreaked onto new plates using quadrant streaking, shown in **Figure 3.1**; the plates were left to culture in anaerobic conditions for up to 2 days. The process was repeated 3 times to isolate the chosen bacterial strain.

The isolated bacterial strains were inoculated in 10 ml of supplemented BHI media broth and left to culture in anaerobic conditions for up to 3 days. The cultures were centrifuged at 4,000 RPM, using the Eppendorf Centrifuge 5810 R, for 15 minutes and the supernatant was disposed of. The pellets were resuspended in 1ml of BHI + 30% glycerol with 200 μl then transferred to a lysing matrix E tube ready for sequencing and the remaining solution was deposited into a cryovial to be stored at -80 °C for long-term storage.

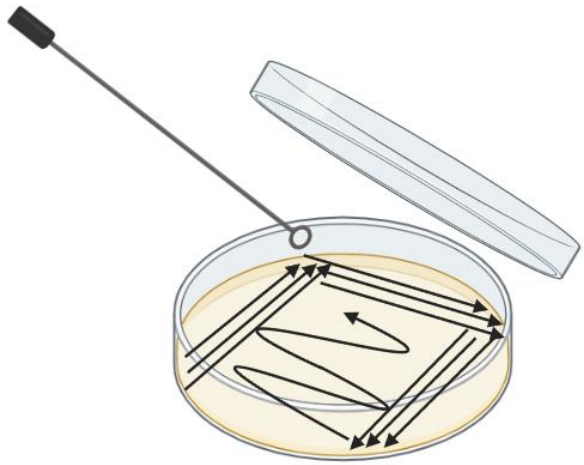

Figure 3.1: Graphical representation of the quadrant streaking technique which allows the formation of isolated colonies on an agar plate.

3.1.2 DNA extraction of faecal samples

The Promega Maxwell RSC PureFood GMO and Authentication Kit was used for the isolated bacterial samples. Approximately 200 μg of the bacterial stock was homogenised with 1ml CTAB in a MP Biomedicals Germany GmbH lysing matrix E tube and vortexed for 30 seconds. The samples were placed on a hot block at 95 °C for 5 minutes before being vortexed for a further minute. The solution was homogenised using the MP Biomedicals Germany GmbH FastPrep instrument for 45 seconds at a speed setting pf 6.0 m/s and 40 μl of proteinase K and 20 μl of RNase A were added and the samples were once again vortexed and placed on a hot block at 70 °C for 10 minutes. Samples were centrifuged at 14,000 x g for 5 minutes to pellet debris and 300 μl of supernatant was added to well 1 of the cartridge. The Maxwell RSC was prepared and then run with the PureFood Protocol following manufacturer's instructions.

3.1.3 PCR and Gel electrophoresis

A 16S rRNA amplification Polymerase Chain Reaction (PCR) protocol was used on the isolated bacterial strains. A master mix was prepared using the reagents shown in **Table 1**. A 96 well transport rack was labelled and 5 μl of the extracted DNA was transferred into the corresponding wells. 45 μl of master mix was added into each well and foil was used to seal the rack. The PCR cycle shown in **Table 2** was run on the Applied Biosystems Veriti Dx Ventiri Thermal cycler.

Table 1: Master mix per reaction for 16s amplification PCR.

10x GC Buffer	10 μl
10 mM dNTP	1 μl
10 uM P1 primer (forward) AGA GTT TGA TCC TGG CTC AG	1 μl
10 uM P2 primer (forward) AGA GTT TGA TCA TGG CTC AG	1 μl
10 uM P2 primer (reverse) ACG GTT ACC TTG TTA CGA CTT	1 μl
H20 (RNA and DNase free)	31.6 μl
Taq polymerase	0.4 μl

Table 2: PCR cycle for 16s amplification.

Start and Taq activation	94 °C for 5 minutes
Denaturing	94 °C for 1 minute
Annealing	43 °C for 1 minute
Extension (35 cycles)	72 °C for 2 minutes
Final extension	72 °C for 7 minutes
Hold	10 °C

A 1.2% agarose gel was prepared using 1.5 g of agarose, 120 ml of TAE, and 12 μl of Invitrogen SYBR Safe. To separate the samples, 5 μl of the PCR product was mixed with 2 μl of loading dye and placed in the wells of the gel and run at 100 volts for 30 minutes; the DNA was then visualised under UV light.

3.1.4 Qubit Assay reaction and Normalising

The extracted DNA was quantified using the Invitrogen Qubit RNA Broad range (BR) assay kit with a visual overview of how to set up the assay reaction shown in **Figure 3.2**. The samples were the vortexed for 2-3 seconds and incubated at room temperature for 2 minutes. The DNA was then quantified with a Qubit 2.0 Fluorometer, and the original sample concentration was calculated.

Using the results identified from the qubit assay, the samples were normalised using RNA and DNase free water to be sent for WGS.

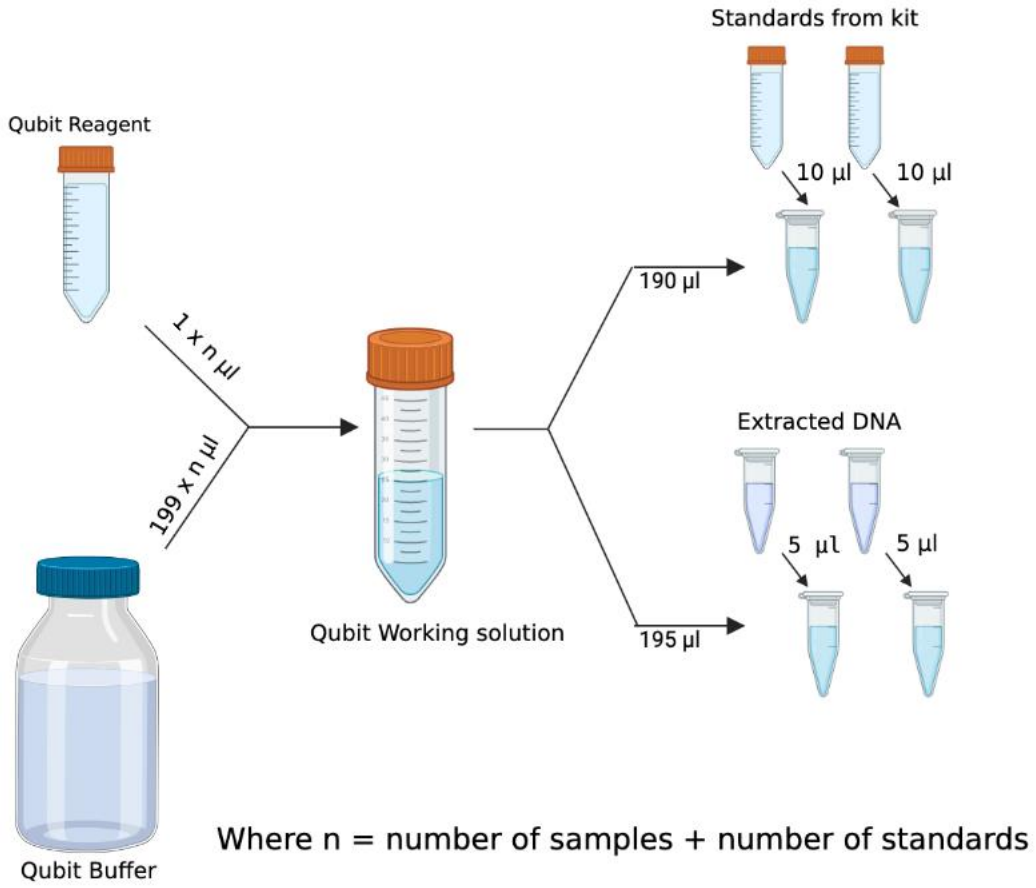

Figure 3.2: A graphical overview of the qubit assay reaction set up to quantify DNA.

3.1.5 Whole Genome Sequencing

The genomes were sequenced using an Illumina NextSeq500 platform with read lengths of 125 bp (paired-end reads); this was done by David Baker and Rhiannon Evans in the internal sequencing team at QIB. The novel technique uses transposase-based library preparation of ARTIC PCR products, (114).

3.2.1 Bacterial species identification

The genomes were first run through fastp-0.20.0. (115) to filter the raw reads to remove low-quality reads. Both Spades-3.11 (116) and the –careful flag. were used for *de novo* assembly and a script written by Dr Raymond Kiu filtered out anything less than 500bp. The genomes were then run through a second script, also written by Dr Raymond Kiu, called BactSpeciesID, to ensure there was no contamination; which worked by extracting 16S rRNA sequences using ABRicate(v1.0.1), specifically using the SILVA-16S database to extract and compare. The script then uses blastn on the genomes to see if there is a match, the default is 99%, to identify

the species of the isolates; the first 8 isolates were initially identified by another member of the lab. This full process was run on the command line.

The genomes were then run through GTDBTK-1.5.1 (117) to later confirm the identification of the species were accurate. GTDBTK uses the Genome Database Taxonomy GTDB and identifies various bacterial and archaeal marker genes using HMMER on whole genome sequences compared with BactSpeciesID which just looks at the 16S rRNA sequences. Genome domains are identified by the highest proportion of marker genes located and placed into the GTDB reference tree using pplacer. The taxonomic classification uses the placement in the tree in combination with relative evolutionary divergence (RED) and ANI for establishing taxonomic ranks.

3.2.2 Phylogenetic tree creation

The genomes were run through Prokka-1.14.6 (118) to produce annotated gff files in GFF3 format. The gff files were then run through Roary-3.12.0 (119) to produce the pan genome and identify the core and accessory genes. The core gene alignment file was produced using snp-sites-2.3.3 (120) which was then run through iqtree-2.0.5 (121) to produce the final newick tree file that was uploaded to iTOL (122) for annotation.

3.2.3 Human Milk Oligosaccharide catabolic clusters analysis

Abricate-1.0.1 was run on the genome sequences using a Human Milk Oligosaccharide database curated by Dr Raymond Kiu which screens for known contigs to identify the presence of genes of interest in 6 different clusters. The isolates were coded by fully present, if all genes of interest of a cluster were present, partially present, if at least one gene of interest in the cluster were present, or absent, if no genes in the cluster were present. HMO data was plotted using the ComplexHeatmap package version 2.10.0 (123) in RStudio version 2022.7.1.554 with R version 4.1.2. Individual isolates and HMO cluster occurrence within isolates were clustered by calculating a Bray-Curtis dissimilarity matrix using the verdict function in the vegan package version 2.6-2 and then hierarchical clustering using the hclust function. The clustered data were plotted as a heatmap using ComplexHeatmap with isolates as rows and HMO gene clusters as columns and species indicated by a colour coded sidebar to denote full presence, partial presence, or absence; this was completed by Dr Matthew Dalby.

3.2.4 Antimicrobial Resistance gene identification

Antimicrobial Resistance gene identification

The sequences were run through Abricate-1.0.1 using the pre-downloaded database Resfinder which identifies genes and/or chromosomal mutations that confer antimicrobial resistance.

4.0 Covid Statement

The amount of laboratory and computational work, and therefore results obtained, has been greatly impacted by COVID-19 regulations, and individual illness. Thus, the original scope of the MRes has been slightly altered and refocused on just the maternal microbiome.

5.0 Results

5.1.1 Participant information

The samples used during this project were a subset derived from the ethically approved and Quadram-BBSRC funded The PEARL cohort. The overarching aims of this longitudinal study are to further our understanding of how microbial diversity and functional capacity changes during the course of pregnancy, the processes of vertical transmission between mother to baby, and how particular microbes interact and affect early life health. Overall, the study recruited 250 pregnant women and their new-born's, and collected, stool, blood, urine, breast milk, skin swabs and low vaginal swabs during pregnancy, birth, and up to 24 months following birth. For this study, 12/250 mothers were selected and further details on these participants are shown in **Table 3:** Participant information**.**

Table 3: Participant information

Participant ID	**Age**	**Antibiotics used previously**	**Probiotics used previously**	**Diet**	**Living with pets**
E007	38	No	No	Both plants & meat	Yes
E014	34	Within the past year	No	Both plants & meat	Yes
E010	31	Within the past year	No	Both plants & meat	Yes
E030	34	No	Yes	Both plants & meat	No
E002	32	Unknown	Unknown	Unknown	Unknown
E003	32	Within the past year	No	Both plants & meat	Yes
E005	40	No	Yes	Both plants & meat	No
E006	37	No	No	Both plants & meat	Yes
E013	37	No	Yes	Both plants & meat	No
E015	32	Within the past year	No	Both plants & meat	No
E008	29	No	No	Both plants & meat	Yes
E009	42	No	No	Both plants & meat	Yes

5.1.2 Faecal Culturing

Bifidobacterium colonies were cultured (on selective media) and isolated from the PEARL faecal samples (**Table 3**), to allow for further investigation into the species commonly found in gut microbiota of pregnant women.

To optimise the culturing and *Bifidobacterium* isolation process, I conducted a pilot study for the faecal culturing, using one patient, comparing the differences between using raw stool samples or stool samples preserved in 60% glycerol, shown in **Figure 5.1**. The PEARL samples, once aliquoted, were stored in either 60% glycerol or were frozen as raw stool. Although it was hypothesised that the glycerol stocks would produce a higher number of bacteria, there were more raw stool stock vials available, and therefore this pilot study was carried out to determine which sample type was optimal for the project going forward.

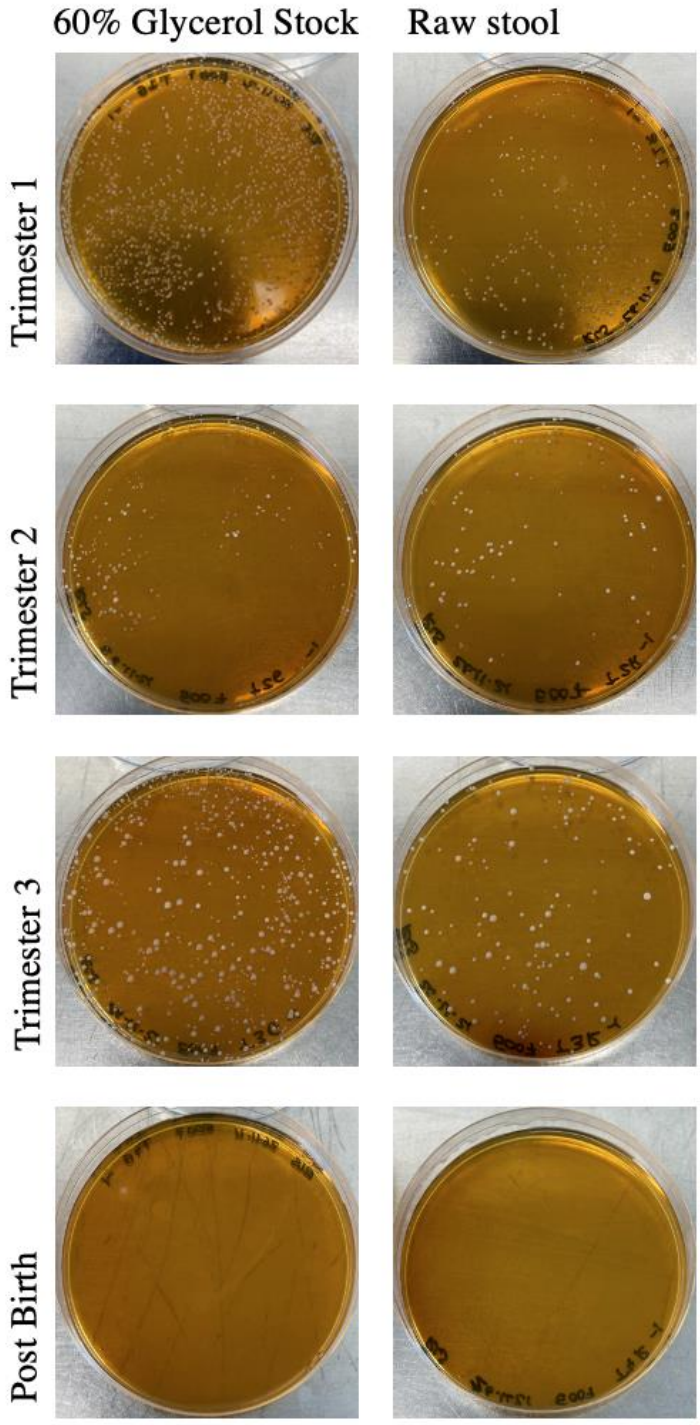

Figure 5.1: MRS agar plates taken after a serial dilution of 10^{-2} and 200 µl of faecal slurry, made up of non-reduced PBS and faecal samples, either using raw stool or stool preserved in 60% glycerol stock from Patient E007, was spread and left to grow in an anaerobic incubator for 2 days.

As expected, more bacteria were present on the plates where the 60% glycerol stock was used, however, as there was still a substantial number of bacteria present on the raw stool plates, I chose to go forward with raw stool samples for the rest of the project as there were more samples/material available. The trimester 3 plates show more bacterial colonies with *Bifidobacterium* morphology (i.e., cream/white, smooth, with a convex circular shape, (124); however, this could not be confirmed until the sequencing data had been received. Only one colony grew on the post birth plate with 60% glycerol stock, and none grew on the plate with raw stool. However, after discussions with the clinical trial coordinator, it was determined these results were likely due to the patient being administered antibiotics post labour.

Non-reduced PBS was used in the pilot study for the dilutions, and I also decided to then test the difference when using reduced PBS to make up the faecal slurry and its effect on bacterial growth (**Figure 5.2)**.

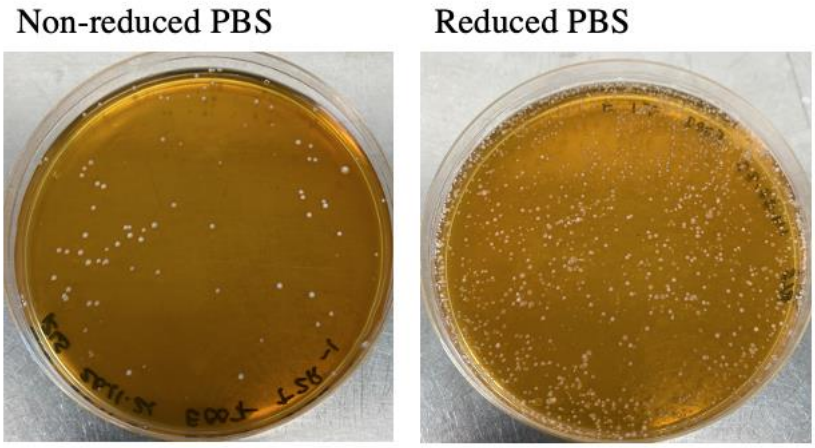

Figure 5.2: MRS Agar plates with 200 µl of faecal slurry, made using non-reduced PBS (L) or reduced PBS (R) spread and incubated anaerobically for 2 days.

The use of reduced PBS had a very strong positive impact on the abundance of bacterial colonies recovered, and thus was used for the rest of the project.

Once the protocol had been optimised, faecal culturing was carried out on a further 11 patients using samples taken from all three trimesters, and one sample post birth, summarised in **Table**

4. Due to the PEARL recruitment protocols, participants can be recruited at any time up to 22 weeks gestation increasing the number of people that can take part in the study, this meant only 3 patients were able to provide samples at trimester one. **Figure 5.3** shows representative agar plates taken from 3 patients at trimester 2/3 and post birth.

Table 4: Overview of the isolated bacteria from faecal culturing

Patient ID	Timepoint	Number of samples isolated
E007	Trimester 1	**9**
	Trimester 2	**3**
	Trimester 3	**3**
E014	Trimester 2	**4**
	Trimester 3	**4**
	Post birth	**5**
E010	Trimester 1	**1**
	Trimester 2	**2**
	Trimester 3	**2**
	Post birth	**4**
E030	Trimester 1	**2**
	Trimester 2	**5**
	Trimester 3	**4**
	Post birth	**5**
E002	Trimester 2	**5**
	Trimester 3	**5**
	Post birth	**5**
E003	Trimester 2	**3**
	Trimester 3	**2**
	Post birth	**5**
E005	Trimester 2	**4**
	Trimester 3	**5**
	Post birth	**4**
E006	Trimester 2	**5**
	Trimester 3	**5**
	Post birth	**4**
E013	Trimester 2	**5**
	Trimester 3	**5**
	Post birth	**5**
E015	Trimester 2	**5**
	Trimester 3	**5**
	Post birth	**5**
E008	Trimester 2	**5**
	Trimester 3	**5**
E009	Trimester 2	**5**
	Trimester 3	**4**
	Post birth	**5**
Total		**159**

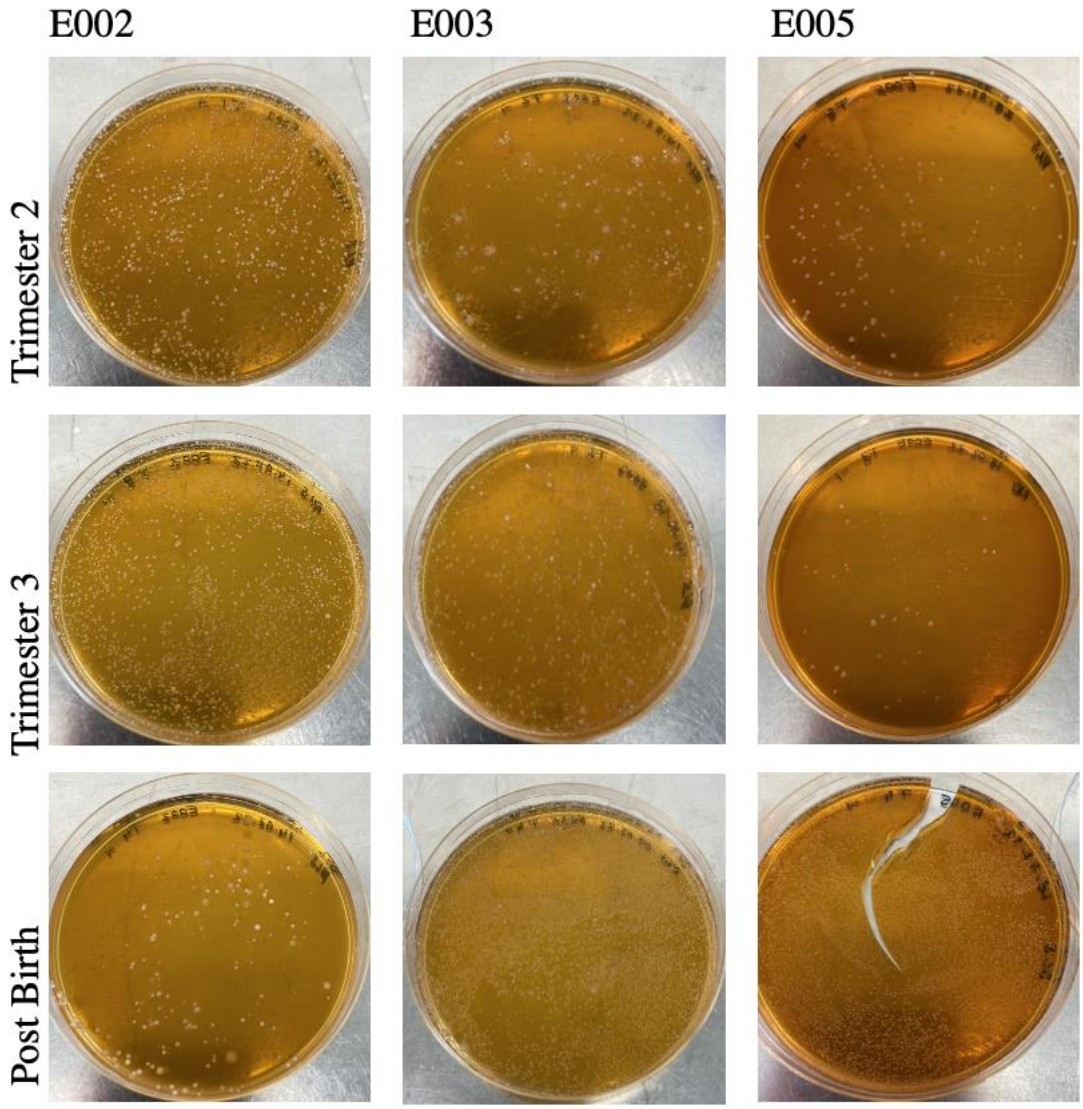

Figure 5.3: MRS Agar plates with 200 µl of faecal slurry spread, made using reduced PBS and raw stool, from three separate patients, taken at three timepoints during pregnancy and post birth. All post birth samples were taken a maximum of one week post labour.

At trimester 2, several bacterial colonies with differing morphologies were observed for both patients E002 and E003, with E003 demonstrating the most diversity. For patient E005 there appeared to be less diverse bacterial colonies, with most looking rounded and light in colour. The abundance of bacterial colonies which grew increased at trimester 3 for both E002 and E003, but reduced for E005 (potentially due to a variety of factors e.g., antibiotics, change in diet etc.). More bacterial colonies with *Bifidobacterium* morphology can be seen for all three participants at post-birth; participant E002 had less bacterial colonies at this time-point, although many of these colonies appeared to have *Bifidobacterium*-like morphology. For both

E003 and E005 there appeared to be an increase in the number of bacterial colonies, with both samples displaying similar results.

5.1.3 Gel Electrophoresis

Initially, DNA was extracted from the bacterial isolate stocks to allow presumptive identification using a full length 16S rRNA PCR protocol, before sending to sequencing. Gel electrophoresis was used to confirm that the PCR was successful, however at this stage the results were not always as expected (**Figure 5.4/5)**, and further details are explained in *6.1.2 16S Sequencing Trouble shooting*.

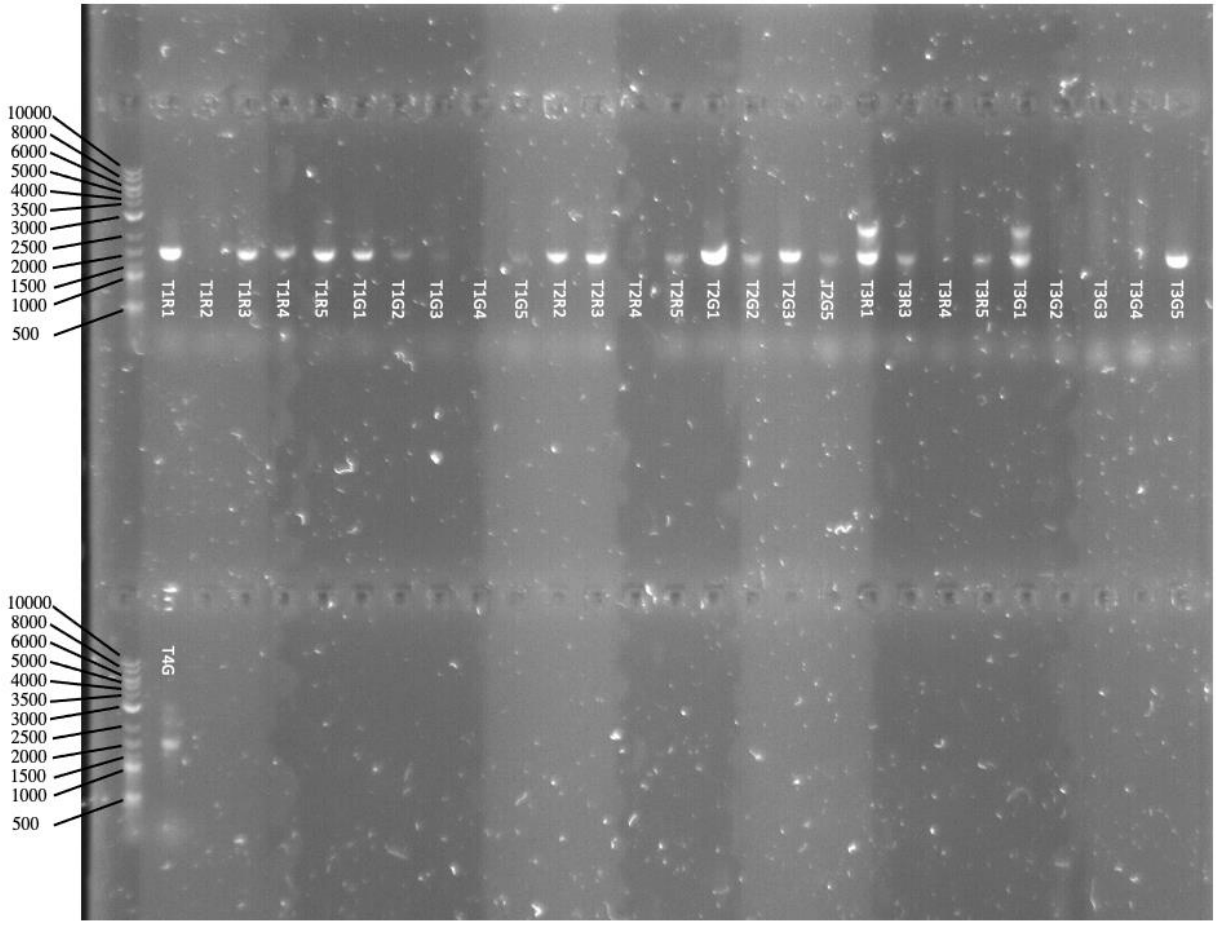

Figure 5.4: Gel electrophoresis (1.2% Agarose) of extracted DNA from Patient E007 in the pilot study.

In **Figure 5.4,** 18/27 lanes show bands at around 1.5kbp indicating the PCR reaction had worked, however 7 of these had a weak signal and both samples E007 T3R1 and T3G1 displayed dimerisation, which could be due to contamination in the sample or human error

when pipetting. Nine of the isolates did not produce a band, which suggested there may be issues with the PCR protocol.

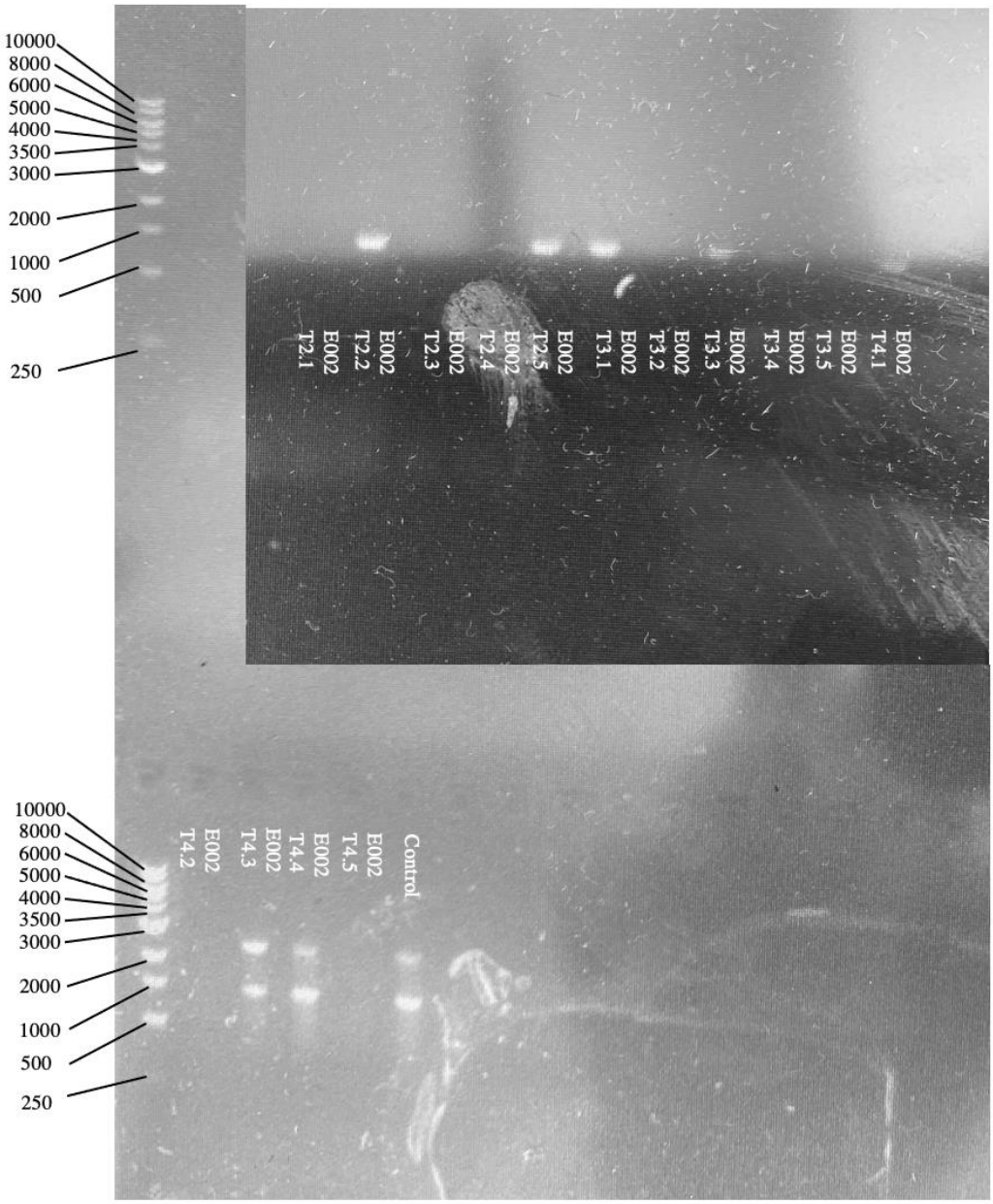

Figure 5.5: Gel electrophoresis (1.2% Agarose) of extracted DNA from Patient E002.

For bacterial isolates from E002, 8/15 bacterial isolates produced an expected band of around 1.5kbp, however 2 of the isolates, E002 T4.3, T4.4, and the positive control revealed dimerisation implying the presence of contamination. Indeed, as the positive control had dimerisation, it can be concluded this was most likely due to human error rather than contamination in the original sample. Seven of the 15 isolates did not produce a band, which

was consistent with the other isolates taken from other patients, again highlighting issues with the PCR protocol.

5.1.4 16S rRNA Sequencing

Due to the inconsistent results observed from the gel electrophoresis, only a selection of samples were sent off externally for 16S rRNA gene sequencing to allow for the initial bacterial identification, all of which had strong bands on the gel indicating the PCR had worked successfully. Unfortunately, all of samples sent off failed QC due to very low-quality reads. Given, I observed strong bands in the gel electrophoresis, the poor-quality data suggests that there may have been an issue during transportation of the samples.

5.2.1 General genomic features

Due to the ongoing issues with initial bacterial identification using 16S rRNA PCR, I decided to move forward with WGS in order to expedite data for downstream analysis. Thus, all samples were sent off to the QIB sequencing team led by David Baker and supported by Rhiannon Evans, whom completed the pre-QC check on all isolates before sequencing. Genomes were assembled and filtered, using fastp-0.20.0, Spades-3.11 and the –careful flag, (115, 116). Using Sequence-stats-1.0, written by Dr Raymond Kiu, *Bifidobacterium* isolate genomic statistics were formulated and compared. Genome sizes ranged from 1.92 Mb (E003-T3.4) to 2.72 Mb (E010-T4.1), with contigs ranging from 24 (E014-T3.5) to 118 (E010-T3.1) per isolate, (**Table S.1**). Two isolates had high values for both genome size and contigs suggesting contamination and were therefore removed from downstream analysis. G+C content ranged from 58.36% (E010-T3.2) to 63.01% (E030-T4.3) (**Table S.1**), with an average of 58.36% which is in line with previous research on *Bifidobacterium*, (125).

5.2.2 Bacterial species identification

Isolate genomes were then run through several programs, including BactSpeciesID and then GTDBTK-1.5.1, (117), to confirm species identity. BactSpeciesID works by comparing the 16S rRNA sequences to those found on the BLAST database to find the most similar match, while GTDBTK uses the GTDB to identify various bacterial and archaeal marker genes in whole genomes and is therefore seen as more reliable. The species (and as appropriate subspecies) identified are summarised in Table 5. Both databases agreed on species identification of every isolate excluding one, further details on this can be found in *6.2.3 Phylogenetic trees and anti-microbial resistance genes.*

Table 5: Summary of identified bacterial species isolated from faecal culturing.

Bacterial Species & subspecies	**Number of isolates identified**	
Bifidobacterium longum subsp. *longum*	Trimester 1	1
	Trimester 2	20
	Trimester 3	26
	Post-birth	16
	Total	**63**
Bifidobacterium animalis subsp. *lactis*	Trimester 1	9
	Trimester 2	17
	Trimester 3	15
	Post-birth	16
	Total	**57**
Bifidobacterium adolescentis	Trimester 1	1
	Trimester 2	4
	Trimester 3	3
	Post-birth	4
	Total	**12**
Bifidobacterium dentium	Trimester 1	-
	Trimester 2	1
	Trimester 3	2
	Post-birth	2
	Total	**4**
Bifidobacterium bifidum	Trimester 1	1
	Trimester 2	2
	Trimester 3	-
	Post-birth	1
	Total	**4**
Eubacterium limosum	Trimester 1	-
	Trimester 2	-
	Trimester 3	1
	Post-birth	-
	Total	**1**

Clostridium perfringens	Trimester 1	-
	Trimester 2	7
	Trimester 3	3
	Post-birth	-
	Total	**10**
Clostridium butyricum	Trimester 1	-
	Trimester 2	-
	Trimester 3	-
	Post-birth	7
	Total	**7**
Paraclostridium bifermentans	Trimester 1	-
	Trimester 2	-
	Trimester 3	-
	Post-birth	1
	Total	**1**

Out of all 159 isolates, 88% isolated were from the *Bifidobacterium* genus; 140 isolates across five different species: *Bifidobacterium longum*, *Bifidobacterium animalis*, *Bifidobacterium adolescentis*, *Bifidobacterium dentium*, and *Bifidobacterium bifidum*; with the majority identified as either *Bifidobacterium longum* or *Bifidobacterium animalis*.

5.2.3 Phylogenetic trees and anti-microbial resistance genes

Once the genomes were assembled, those which had been identified as belonging to the *Bifidobacterium* genus were run through Prokka (118) and Roary (119) pipelines to generate the pangenome for the selected isolates, the summary statistics are shown in **Table 6**. The Single Nucleotide Polymorphisms (SNP)s were extracted using SNP-Sites-1.0 (120) to produce a TREEFILE, coupled with reference genomes (downloaded from BLAST and corresponding to the species identified in this study), accession numbers for the reference genomes are listed in **Table 7**. The TREEFILE was then uploaded to iTol (122) to create a phylogenetic tree of all the isolates, shown in **Figure 5.6**. The phylogenetic tree was used to visualise the phylogenetic relationships of the isolates between and within the different species. The 5 distinct *Bifidobacterium* species can be clearly identified with many of the isolates from

the same participant clustering together in clades; the *B. animalis* isolates however did not have distinct clades, but were rather distributed throughout the tree.

The presence of antimicrobial resistance genes (ARGs) encoded within the human gut microbiome represents a threat to health. Moreover, there has been previous work suggesting that multi-drug resistant bacteria and ARGs identified in the maternal microbiome, may be passed onto offspring through vertical or HGT. Thus, the selected genomes were also run through ABRicate (126) using the database ResFinder to search for three specific ARG's: *aph(3')-Ia_7* which aids neomycin and structurally related aminoglycosides, the erythromycin resistance gene *erm*(X) commonly found in *Bifidobacterium*, and *tet(W)_4* which confers resistance to tetracycline, (127-129). Those that could be identified were annotated onto the phylogenetic tree using blue bars which can be seen in **Figure 5.6**.

Table 6: Summary statistics for the pangenome.

% of genome	**Number of genes**
Core genes (99% <= strains <= 100%)	22
Soft core genes (95% <= strains < 99%)	0
Shell genes (15% <= strains < 95%)	3553
Cloud genes (0% <= strains < 15%)	16412
Total genes (0% <= strains <= 100%)	19987

Table 7: The accession numbers for the reference genomes used in the roary analysis.

Species	**Accession numbers**
Bifidobacterium animalis	GCA_000022705.1
	GCA_000022965.1
	GCA_000025245.2
	GCA_000092765.1
	GCA_000220885.1
	GCA_000260715.1
	GCA_000277325.1
	GCA_000277345.1
	GCA_000414215.1
	GCA_000471945.1
	GCA_000695895.1
	GCA_000020425.1

	GCA_000471945.1
Bifidobacterium adolescentis	GCA_000737885.1
	GCA_009832825.1
	GCA_000817995.1
	GCA_017815835.1
	GCA_003856735.1
	GCA_003429385.1
	GCA_003030905.1
	GCA_902386735.1
	GCA_000010425.1
Bifidobacterium bifidum	GCA_002845845.1
	GCA_020892075.1
	GCA_000265095.1
	GCA_003390735.1
	GCA_001281345.1
	GCA_003573895.1
	GCA_000466525.1
Bifidobacterium dentium	GCF_017743195.1
	GCF_900637175.1
	GCF_000024445.1
	GCF_017743215.1
	GCA_900637175.1
Bifidobacterium longum	GCA_000020425.1
	GCA_021228035.1
	GCA_014898115.1
	GCA_000092325.1
	GCA_000730205.1
	GCA_001293145.1
	GCA_000196555.1
	GCA_023205815.1
	GCA_023208115.1
	GCA_014898135.1
	GCA_000196575.1

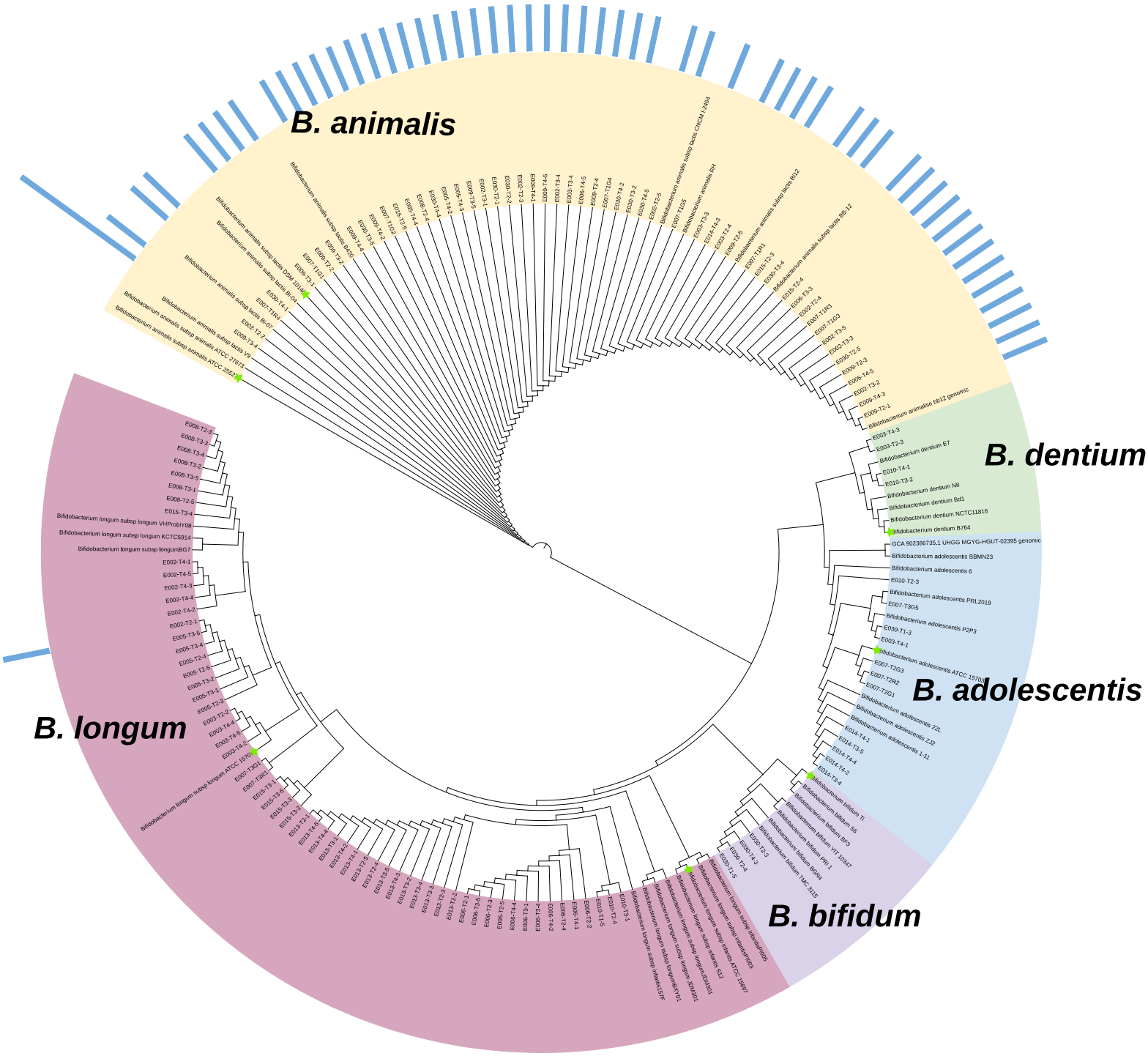

Figure 5.6: Phylogenetic tree showing the evolutionary relationships between different bacterial species which were isolated from faecal samples of pregnant women. Reference genomes have been included in the analysis with the type strains for several species/sub-species annotated with a green star. ARG data is highlighted via blue bars next to the isolate ID.

Fifty-six out of 159 isolates contained the ARG *tet(W)_4* , with isolate E009-T3-4 being the only isolate to have all 3 ARG identified in its genome. Notably, fifty-five out of 56 of these

isolates belonged to the *B. animalis* species with one isolate, E002-T2-1, identified as *B. longum*.

Further analysis was carried out on the two species for which I had the most isolates: *B. longum* and *B. animalis*. Phylogenetic trees were created which were species specific to investigate how the isolates were related at species level and to identify sub-species patterns; this can be seen in **Figure 5.7** and **Figure 5.8**; I also identified clonal strains using SNP distances, calculated for both species **(Figures S2** & **S3)**, with values <2 assumed as clonal.

In **Figure 5.7**, which represents *B. longum***,** many of the clades are split into isolates belonging to a single participant/ and or timepoint. Participant E002 had clustering of 5 isolates at the post-birth timepoint, however SNP distances indicated these were not clonal. All isolates taken from participant E005, at both trimester 2 and 3, also clustered together with low SNP distances observed (**Table S.2**); 2 out of 7 isolates (E005-T2.5 & E005-3.2) had SNP distances < 2 suggesting clonal strains, even though they were taken from separate timepoints. Six out of seven of the E005 isolates had SNP distances <8 suggesting long-term colonisation. Similar results were also seen for isolates from participants E006 and E013, however much larger values were observed for SNP distances. Several isolates were shown to lie separately on the tree from the other isolates taken from the same participant, including E002-T2.1, E008-T2.2 and, E015-T3.4, all of which had SNP distances that were > 7,000 suggesting they are different strains from the other isolates taken from the same participant. FastANI analysis identified all 64 isolates to be *Bifidobacterium longum* subsp. *longum*.

Contrasting with the results above, **Figure 5.8** (showing *B. animalis*) indicates that there were no distinct clades between timepoints nor between the different participants, excluding the separation of the two reference genomes *B. animalis* subsp. *animalis ATCC27673* and type strain *B. animalis* subsp. *animalis ATCC2552*. These results link to those seen in **Table S.3,** as all isolates, and *B. animalis subsp. lactis* reference genomes, displayed very low values for the calculated SNP distances many of which were < 40. The FastANI analysis also revealed all isolates to belong to *B. animalis* subsp. *lactis*. No patterns could therefore be determined from these results, rather these data indicates potential common sources for this subspecies.

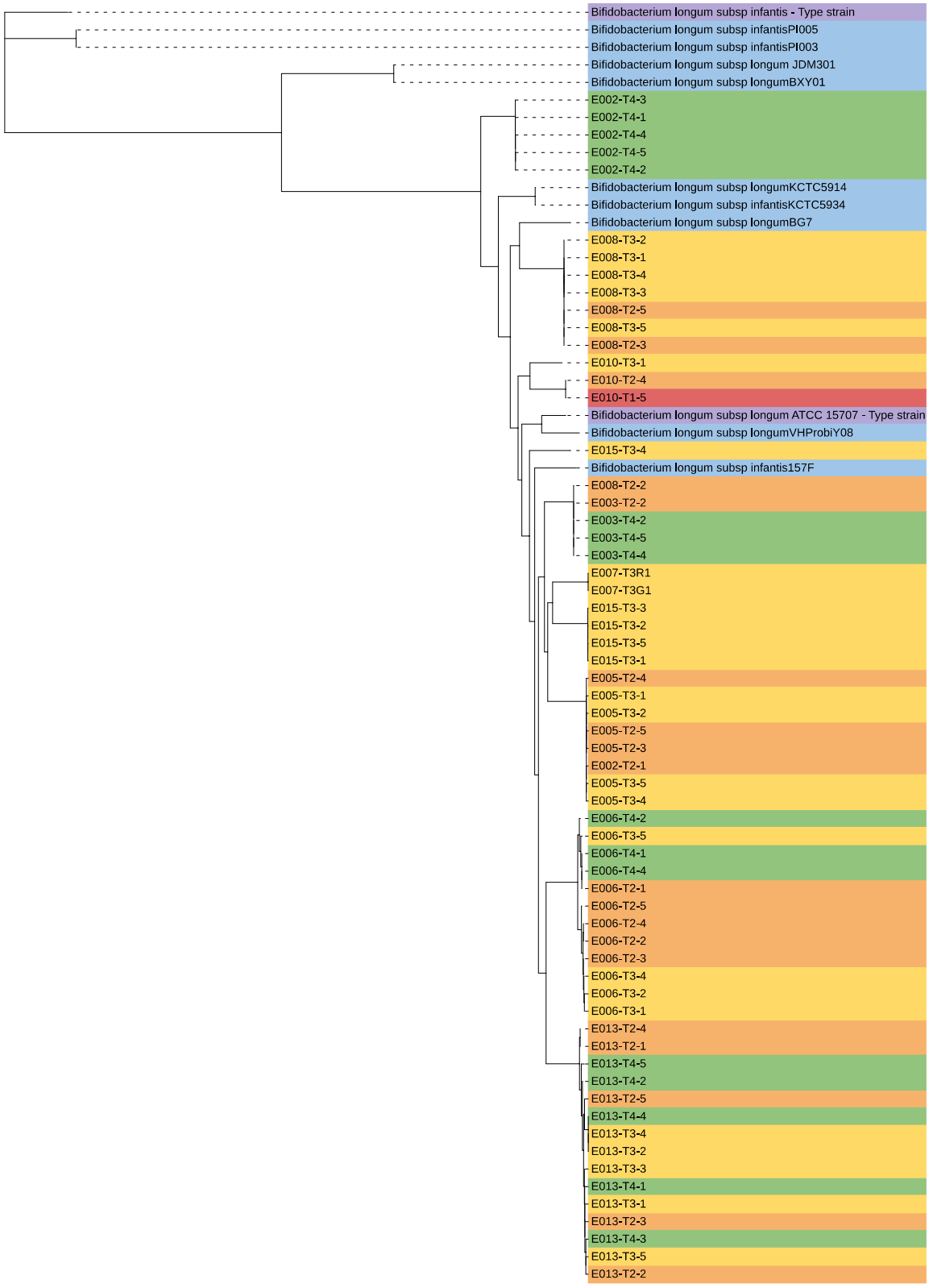

Figure 5.7: Phylogenetic tree visualising the evolutionary relationships between 63 bacterial isolates identified as *Bifidobacterium longum* cultured from faecal samples taken throughout pregnancy and post birth; including 10 reference strains with the type strains for *Bifidobacterium longum* subsp. *longum* and *Bifidobacterium longum* subsp. *infantis*. Samples were isolated from 9/12 pregnant

participants and are colour coded depending on the timepoint the faecal samples were taken from: with one isolate from trimester 1 shown in red, 19 isolates from trimester 2 shown in orange, 26 isolates from trimester 3 shown in yellow and 16 isolates taken post-birth (up to week 1 post-birth) shown in green.

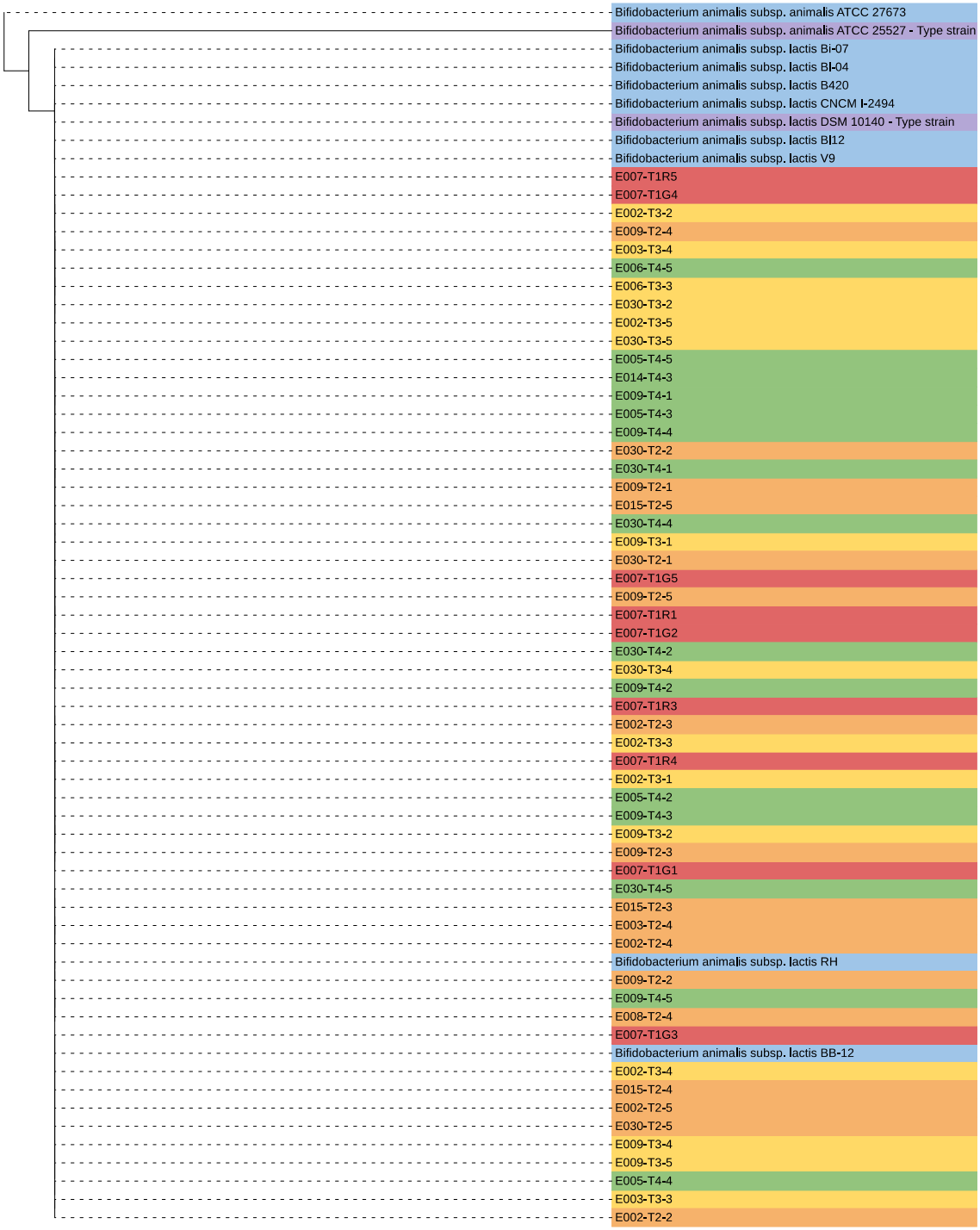

Figure 5.8: Phylogenetic tree of 57 isolates which were identified as *Bifidobacterium animalis*, isolated and cultured from 10/12 maternal participants at different stages throughout pregnancy and post-birth.

Isolates have been colour coded depending on time-point: 9 isolates were taken from stool samples during trimester 1 (red), 17 isolates were taken during trimester 2 (orange), 15 isolates were taken during trimester 3 (yellow), and 14 isolates were taken post-birth (green).

5.2.4 Human milk oligosaccharide clusters

One important function associated with many *Bifidobacterium* species, particularly those found within the early-life microbiota, are their ability to target and break down HMOs (130). These genes which are specialised for the degradation of HMOs, are found in HMO gene clusters, and are extremely valuable for infant health.

The genomes of isolates which had been identified as belonging to the *Bifidobacterium* genus were run through Abricate-1.0.1 using a curated database to look for 6 known HMO clusters. The resulting data was visualised using the ComplexHeatmap package version 2.10.0 (123) in RStudio version 2022.7.1.554; which can be seen in **Figure 5.9**. Isolates belonging to the species *B. longum* were identified to contain the highest percentage of HMO clusters in their genomes, with every isolate having the full HMO cluster B_breve_UCC2003_BBR_RS18490-BBR_RS18520 present. The BBR_RS18490-BBR_RS18520 cluster, otherwise known as the *nah* cluster, has previously been described in *B. breve* UCC2003, along with the *lnt* cluster, (BBR_RS13075-BBR_RS13100), *lac* cluster (BBR_RS18470-BBR_RS18480), and the *lnp/glt* cluster (BBR_RS18650-BBR_RS18675), (131), all of which were identified in several of the isolates. The *lac* cluster was only fully present in the four *B. bifidum* isolates, although it was found partially present in all *B. animalis* isolates and seven *B. longum* isolates. Overall, the *B. animalis* isolates contained the lowest percentage of HMO clusters in their genomes, none of which were found fully present, with only the *lac* and *lnt* clusters found partially present. The *B. dentium* isolates also did not contain any complete HMO clusters, however, the *lnt* and *nah* clusters were found partially present as was the 45 kb HMO cluster originally found in *B. longum* subsp. *infantis* (BLON_RS12095_RS122215) which is involved in the digestion of several HMOs, (132). Only 8 isolates were found to contain all 6 HMO clusters at least partially present (E008-T2.5, E008-T3.1, E008-T3.2, E008-T3.3, E008-T3.4, E008-T3.5, E010-T1.5 E010-T2.4), all of which were *B. longum* isolates.

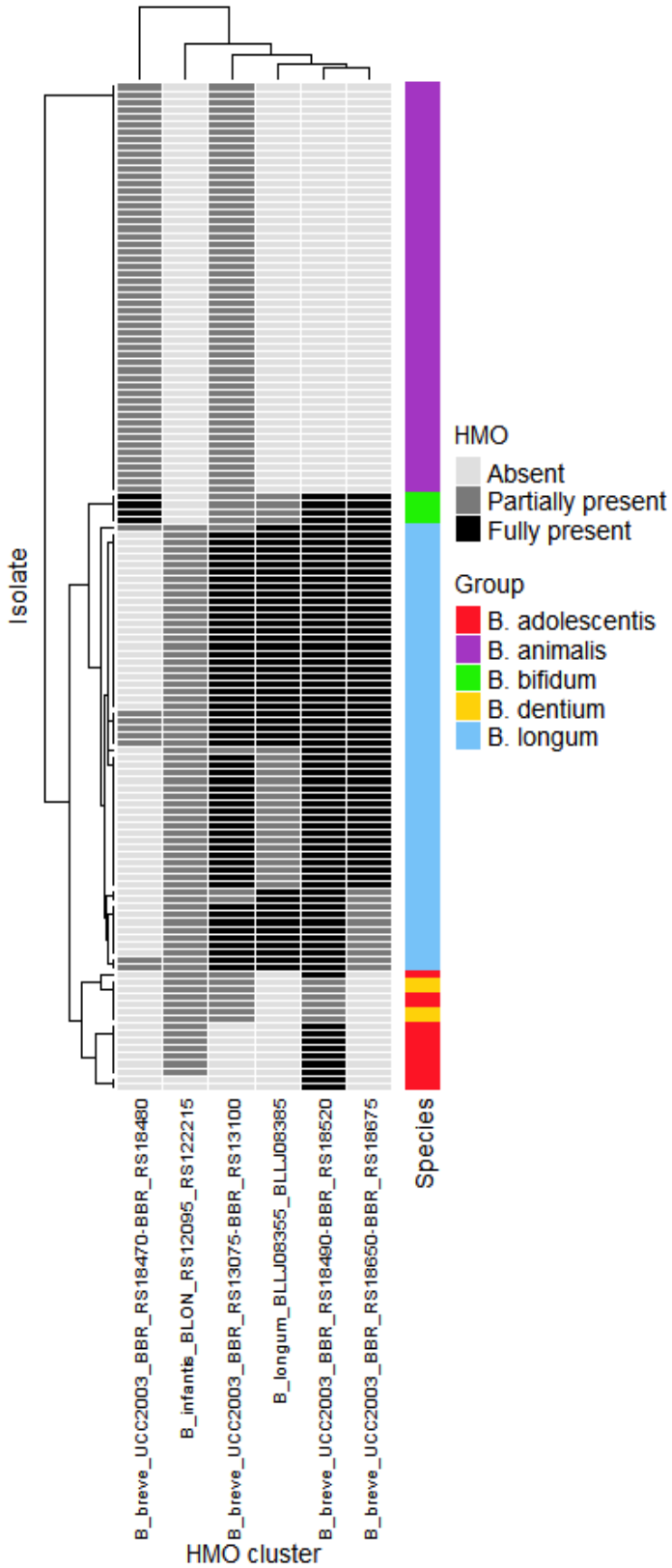

Figure 5.9: A heatmap plotted using the ComplexHeatmap package version 2.10.0 in RStudio which illustrates the presence of 6 known HMO clusters found in the genomes of *Bifidobacterium* isolates which were cultured from pregnant participants throughout gestation.

B_breve_UCC2003_BBR_RS13075-BBR_RS13100 catabolises lacto-*N*-tetraose, B_breve_UCC2003_BBR_RS18470-BBR_RS18480 catabolises lactose, B_breve_UCC2003_BBR_RS18490-BBR_RS18520 catabolises *N*-acetylhexosamine, and B_breve_UCC2003_BBR_RS18650-BBR_RS18675 catabolises lacto-*N*-biose phosphorylase. B_infantis_BLON_RS12095_RS122215 and B_longum_BLLJ08355_BLLJ08385 can catabolise multiple HMOs including lacto-*N*-tetraose, lacto-*N*-biose and lactose. Isolates were categorised as fully present, black, if all genes of the cluster could be found in the genome, partially present, dark grey, if 1 or more were identified, or absent, light grey, if none were found. The species of each isolate was added below the heatmap. This was created by Dr Matthew Dalby.

6.0 Discussion

The potential importance of the maternal microbiome, for both mother and infant health, has been gaining interest; however, our understanding of its composition and the changes it undergoes throughout the gestation period remains under-researched. This is particularly true for beneficial genera that are associated with improved host health such as *Bifidobacterium*. Thus, with a focus on the microbiota genus *Bifidobacterium,* this project aimed to identify and compare which species and strains reside in the maternal gut microbiota throughout pregnancy and at birth using both culturomics and genomic analysis approaches. A total of 140 *Bifidobacterium* colonies were isolated from 12 pregnant women that represented 5 species, with the majority identified belonging to *Bifidobacterium longum* and *Bifidobacterium animalis*. In some cases, I observed carriage of strains within individuals over the course of pregnancy (e.g. *B. longum*), whereas for others there did not appear to be any obvious intra-individual patterns (e.g. *B. animalis*). These strains encoded a limited number of ARG (almost exclusively within *B. animalis*), and HMO clusters were found at least partially present in all 5 species, with those found fully present in 3/5 species.

These results offer insights into the diversity of bifidobacteria species and strains associated with the maternal microbiota. Investigating the genomes of these isolates may provide an explanation as to why these species/strains are favoured over others, and how they contribute towards maternal and consequently offspring health and metabolism.

6.1.1 Faecal culturing

The pilot study was vital for optimising the protocol for faecal culturing. I first investigated the viability of using stool samples that had been preserved in 60% glycerol solution and frozen, compared to raw stool samples that had been frozen; these results can be seen in **Figure 5.1**. As mentioned in *5.1.2 Faecal culturing*, there were a greater number of raw stool aliquots compared to those in 60% glycerol, making it a preferable option for this study. The results demonstrated an increase in viability of bacteria on the plates made with the glycerol solution however, the raw stool plates produced a high enough yield of bacteria for the project and therefore I chose to continue using the raw stool aliquots for the rest of the study. However, it should be noted that the glycerol stocks may better preserve a wider range of *Bifidobacterium* species, and therefore the raw stool sampling may provide a somewhat bias ‘snapshot’ of

overall diversity. Given shotgun metagenomics profiles are also being collected as part of PEARL, this could be checked, and compared back to the WGS data obtained from this pilot experiment to determine if there are any differences.

Both plates taken from the post-birth timepoint had a distinct lack of bacteria, with only one colony observed on the 60% glycerol plate. However, participant medical records showed the patient was administered antibiotics at birth which may explain this result. Routine antibiotic treatment is often administered to women who have given birth to reduce the risk of maternal infections (i.e. routine antibiotic prophylaxis) (133), however this could have a negative effect on the mother's gut microbiota. Chen et al investigated the effects of antibiotics on the maternal microbiota in mice and found the composition of the gut bacteria was significantly affected, upon the administration of antibiotics, as well as an increased intestinal injury score and cytokine levels were observed, (134). One systemic review also correlated data from several clinical trials and reported the use of antibiotics did not reduce the risk of urinary tract infections, wound infection or the duration of hospital length, (133), suggesting the use of routine antibiotics may not be as effective as previously believed. My results show a loss in bacterial abundance post-birth for this patient, also demonstrating how gut bacteria may be affected from antibiotic use. Subsequently, I later discovered 15/15 of the isolates from the pilot study were in fact *Bifidobacterium*, of which one strain can be seen in timepoints trimester 2 and 3, highlighting how the use of antibiotics can impact the abundance of both beneficial and pathogenic bacteria in the gut. As shown in my results, *5.2.3 Phylogenetic trees and anti-microbial resistance genes*, *Bifidobacterium* are not multi-drug resistant making them particularly susceptible to many antibiotics; adding to the issues surrounding administration of antibiotics during pregnancy. Considering the association of maternal antibiotic use, during pregnancy and post-labour, with adverse infant health outcomes, (135), these results (albeit on a very small number of samples), coupled with previous studies, indicates that careful consideration of prophylactic antibiotic use in mothers should be undertaken.

During the pilot study I also investigated the difference in bacterial growth when using reduced PBS compared to non-reduced PBS when making the faecal slurries. As expected, reduced PBS produced a far greater yield of bacteria compared to the plates made up with non-reduced PBS as shown in **Figure 5.2**, as bifidobacteria are strict anaerobes, (136).

For each participant, I isolated 5 bacterial colonies from each timepoint producing a maximum of 10-20 isolates for all 12 patients dependent on how many timepoints they provided samples for. **Table 4** summarises the number of bacterial isolates that were obtained during the project. Previous studies have demonstrated infants contain multi-*Bifidobacterium* strains in the gut microbiota and therefore several colonies were picked from each plate to investigate the possibility of multi-strains existing in the maternal microbiota, (137). In total, 159 bacterial colonies were isolated and sequenced from a total of 185 possible isolates. Unfortunately, some of the bacterial colonies were lost due to issues with the initial experimental methodology; re-culturing isolates growing on MRS agar plates into MRS liquid culture to produce bacterial stocks. After observing this issue, I chose to grow the isolates in BHI liquid media, once they had been isolated on the MRS agar plates, which subsequently prevented the large loss in cultures. Different *Bifidobacterium* species show a preference to different media types, (138), and therefore it is possible that the cultures that were lost do not grow well in MRS liquid media; these results show the importance in optimising culturomics approaches to gain a greater understanding into the requirements and characteristics of different bacterial species and strains.

There appeared to be a distinct difference in bacterial diversity and abundance on the plates for each participant at differing time-points **(Figure 5.3)**. On each plate there were many colonies which represent the morphology typical to the *Bifidobacterium* genera, (124), with an increase in these types of bacterial colonies seen as the pregnancy progresses. Previous studies have found the abundance of Actinomycetota, including bifidobacteria, increases during trimester 2/3, (44), aligning with what can be seen on the plates. However, it can be difficult to make concluding statements about the microbiota of the participants from the plates alone due to the selective nature of culturing, which may not provide a completely representative sample of overall of particular gut bacteria. Although, selective media was used during the culturing process to produce a greater yield of *Bifidobacterium* isolates, it is known that certain species may need different/additional nutritional requirements which may impact overall profiles obtained; one example of this is the ability of different *Bifidobacterium* species/strains to utilise different carbohydrates, . However, given the labour intensive nature of culturing, a pragmatic approach – as in selecting a '*Bifidobacterium* standard' media was required for this project. There are also issues when using the morphology of bacterial colonies to determine the abundance of bacteria in the gut, and therefore further data was required for these isolates through sequencing to be able to confirm bifidobacteria-like morphology was actually a

Bifidobacterium isolate in reality. If there was more time, Real-time PCR (qPCR) on the original aliquoted sample could provide useful information on the relative abundance of these species, although given how cost-effective WGS is this may not be necessary.

6.1.2 16S Sequencing Trouble shooting

The original plan for this project involved following a full length 16S rRNA PCR protocol after extracting the DNA from the bacterial isolates and consequently sending off for Sanger sequencing to identify which isolates were *Bifidobacterium*, with selected isolates then sent for WGS. Gel electrophoresis was used before the isolates were sent off to confirm to protocol was successful, however there were distinct issues identified at this stage, as seen in **Figures 5.4/5**. Gel electrophoresis was first completed on the DNA extractions from the pilot study, **Figure 5.4**, 7 of the samples had a poor signal and 2 samples demonstrated dimerisation suggesting contamination in the sample, later analysis shows those 2 samples did not have contamination and therefore this result was most likely due to human error when pipetting. This protocol was then completed on the bacterial isolates from patient E002, shown in **Figure 5.5**, and again many of the isolates did not produce a band. Although some of the issues were most likely caused by human error, the gels were repeated and there was still a lack of bands for many of the samples. It was not confirmed what may have been going wrong in the PCR protocol and if I had more time I would investigate this further.

Due to the issues that I had encountered with the PCR protocol I chose to only send a section of the samples for 16S rRNA sequencing externally all of which had produced strong bands on the gel electrophoresis. The results I received from this however were not as expected, with every isolate failing the QC checks due to too low-quality reads.. Two separate external companies were trialled for 16S rRNA sequencing, however both indicated similar results and it was concluded there was a fault during the transportation of these samples to the external companies. This was confirmed after discussions with other members of the laboratory who were having similar issues. For the rest of my study, I opted to send all of my isolates for WGS internally instead, and I did not face any issues with this approach. Given more time I would conduct further research into what was happening in the transportation process which was causing the DNA to degrade however, due to time constraints of this project I was unable to come to a conclusion.

6.2.1 Genome characteristics

As mentioned in *5.2.2 General genomic features*, and seen in **Table S.1**, 138 out of the 140 of the *Bifidobacterium* isolates had genome sizes which aligned with previous publications that is expected from *Bifidobacterium* species, (125). Similarly, the G+C content of the *Bifidobacterium* isolates were also as expected with a mean value of 58.36%, as were the contigs values, matching what has already been seen, (125). Two isolates had contigs values of 1497 and 777 respectively, although their G+C% was in line with the other isolates; these results suggest there may have been contamination in these isolates when they were sequenced which may have been to human error or the original isolates were not pure to start with. These isolates were not considered for further analysis.

6.2.2 Species identification

The 159 genomes were run through two identification programmes, BactSpeciesID and GTDBTK-1.5.1 (117), to identify the isolates at species level. Isolates which had been identified as *B. longum* or *B. animalis* were then compared to different type strains of the species, using FastANI; the results from these analyses are represented by **Table 5**. Out of the 159 isolates, 88% were confirmed to be *Bifidobacterium* belonging to 5 different species: *Bifidobacterium longum*, *Bifidobacterium animalis*, *Bifidobacterium adolescentis*, *Bifidobacterium dentium*, and *Bifidobacterium bifidum* all of which are commonly isolated from the human gastrointestinal tract. The high proportion of *Bifidobacterium* isolates highlights the utility of using a selective media when trying to culture and isolate bifidobacteria; the plates were treated with Mupirocin, due to *Bifidobacterium's* intrinsic resistance, (139). The remaining 12% of isolates were identified as other bacterial species which are also commonly found in the gut: *Eubacterium limosum*, *Clostridium perfringens*, *Clostridium butyricum* and *Paraclostridium bifermentans*; all of which exhibit a similar morphology when grown on an agar plate to bifidobacteria demonstrating why they may have been assumed to be *Bifidobacterium* colonies during the culturing process, (140, 141). Both *Clostridium perfringens* and *Paraclostridium bifermentans* are known to be pathogenic towards humans, with *C. perfringens* being associated with gastroenteritis in adults and the fatal infant disease NEC, (142-145). Due to the risk of bacterial transmission, the presence of *C. perfringens* may be a cause of concern for the offspring health however further investigations provided by the participants health records showed none of the infants were diagnosed with NEC.

The composition of the adult mammalian gut microbiota has been extensively explored, and it is apparent that *Bifidobacterium* is a core constituent of this 'mature' microbial ecosystem . In the adult gut, *B. longum* subsp. *longum* and *B. adolescentis* have been identified to be the most abundant, (146). As mentioned in *1.2.2 The Maternal Microbiota*, the gut microbiome during the first trimester of pregnancy tends to represent what is seen in a typical adult microbiota, therefore I would have expected to see the highest proportion of bacterial colonies, isolated from trimester 1 samples, to be identified as *B. longum* subsp. *longum* and *B. adolescentis.* However, 9/12 of these isolates were identified as *B. animalis* subsp. *lactis*, with only 1 identified as *B. longum* subsp. *longum* and 2 more identified as *B. adolescentis* and *B. bifidum* respectively; although out of the 9 *B. animalis* isolates, 7 of these had SNP distance values >6 suggesting they may be clonal strains (this is discussed further below).

The largest proportion of *Bifidobacterium* isolates belonged to the *B. longum* species, 63/140, with the second largest identified as *B. animalis*, 57/140. As previously stated, the adult gut is typically dominated by *B. longum* and therefore this species being the most abundant in my isolates was expected. *B. longum* are able to metabolise a wide range of carbohydrates and encode a many diverse genes and predicted glycosyl-hydrolases. They are also able to digest both plant derived carbohydrates and HMOs giving a large competitive advantage to their host which may offer insight into why these species are found so abundantly in the human gut, (147). I did not expect to see such a large proportion of *B. animalis* isolates however, these results may coincide with an intake of dairy (146), as *B. animalis* subsp. *lactis* has been identified to be found in human food since 1980, (148), and it is also a commonly used probiotic, (149). However, only 3/12 participants claimed to take probiotics before the study, shown in **Table 3**. In comparison to healthy adults, there is a lack of publications on the composition of the maternal microbiota during the different stages of pregnancy and post-birth, therefore it is difficult to compare the results from this study to what we would expected 'norms'. However, another study conducted by Yang et al did show similar results in the proportion of different *Bifidobacterium* species in the maternal gut, with both *B. longum* subsp. *longum* and *B. animalis* subsp. *lactis* being identified as a higher proportion of the *Bifidobacterium* population, although one distinct difference was the concurrently high proportion of *Bifidobacterium pseudocatenulatum* (150). Previous publications have listed *B. pseudocatenulatum* as abundant in the human gut, (151), and therefore I had expected to see this species within my isolates.

6.2.3 Phylogenetic trees and anti-microbial resistance genes

The phylogenetic relationships between the *Bifidobacterium* isolates and several reference genomes were investigated using a phylogenetic tree created on the iTol platform (122), which is shown in **Figure 5.6**. This analysis displayed a clear distinction across the 5 species with varying levels of clustering.

As the two most abundant species in my dataset were *B. longum* and *B. animalis*, I decided to investigate these two species further, and more details on these isolates are described on page 52.

Twelve isolates from 5 different participants were identified as *B. adolescentis,* shown in the blue section of the phylogenetic tree along with 9 reference genomes and type strain *B. adolescentis* ATCC15703. At the bottom of this clade there is a cluster of 5 isolates from participant E014 taken at timepoints trimester 3 and post-birth. Notably, upon investigating if they were clonal strains, they had SNP distance values >10 indicating that although they are closely related they are not identical strains. *B. adolescentis*, alongside *B. longum*, is one of the most abundant species often found in the adult human gut, (152) and therefore the low numbers observed were somewhat surprising, and may indicate this species is not observed as frequently in pregnant women. *B. adolescentis* has been previously researched in many contexts associated with human health and is often found in a healthy adult gut microbiota, (153), and therefore long-term colonisation may provide a maternal protective effect. Interestingly, 3 of the genomes isolates from participant E007 clustered most closely with the *B. adolescentis* type strain, which were all taken from trimester 2, this strain in particular has been reported to provide a probiotic effect against rotavirus, (154). Originally, one of the isolates had initially been identified as *B. faecale* using BactSpeciesID, however GTDBK analysis identified it as *B. adolescentis*. Upon further investigation into this species, *B. faecale* was originally only published with its 16S rRNA and hsp60 genes, (155) and using FastANI on the published genome, against the *B. adolescentis* reference genome, it received a value >95% showing that this should not be considered a separate species from *B. adolescentis* and should be reclassified. None of the isolates were clonal as they all had SNP values >2.

The phylogenetic relationships of the 4 identified *B. dentium* isolates were investigated alongside 5 reference genomes including the type strain *B. dentium* B764. *B. dentium* is found both in the oral and intestinal microbiome, and has been reported for its ability to adhere to

human MUC2+ mucus and harbour mucin-binding proteins (156). The isolates clustered in 2 separate clades which were participant specific. The two isolates from participant E003 displayed the least genetic relatedness to the reference genomes with SNP distance values all >9880. The two isolates from participant E010 clustered with the reference genome *B. dentium* E7 with SNP distance values of 2912-2930. Although the strains isolated from the same participant were similar, none had SNP distance values lower than 20 and are therefore not considered clonal.

Four isolates from participant E030 were identified as *B. bifidum* which were compared with seven reference genomes, downloaded from the NCBI database, including the type strain *B. bifidum* Ti. *B. bifidum* is typically found in higher abundances in the infant gut rather than the adult gut so the low isolate number was to be expected, (146). The four isolates clustered together with low SNP distance values, the lowest being between E030-T1-5 and E030-T4.3 however, even though the phylogenetic tree shows close clustering between isolates E030-T1-5 and E030-T2-4, this may occur due to there are more genes in common between these two isolates.

A more in-depth analysis was conducted on the isolates identified as *B. longum* shown in **Figure 5.7**. The tree shows several clades which reveal clustering with isolates taken from the same participant. The bottom clade reveals 15 isolates from participant E013 with several of these isolates displaying very low SNP distance values <10 of which cluster together. Three of these smaller clusters include isolates which are present across multiple isolates suggesting there is long term colonisation of these strains in the maternal gut. Similar results was also seen for isolates from E005 and E008 as well as isolates from E006 where clonal strains were observed in two separate timepoints. Many studies have shown the protective ability of *B. longum* in the gut, (157), and therefore it's capability to colonise long term would suggest an evolutionary benefit to its host. All of the isolates were confirmed, using FastANI, to be *B. longum* subsp. *longum* which can be seen on the tree sharing 954 core genes. Several of the *B. longum* subsp. *infantis* reference strains are clustered at the top of the tree, including the type strain however, 2 of the *infantis* strains are shown lower down the tree; further analysis concluded these strains may have been originally misclassified and are actually in fact *B. longum* subsp. *longum*. The infant microbiome is especially dominated by *B. longum* subsp. *infantis* which would suggest bacterial transfer from the mother may play a role in this establishment, (100), however the lack of this sub-species in these isolates suggest another

mechanism could be at play; further analysis into the microbiota of the offspring of these participants is needed to understand where this subspecies is coming from. It is possible there may be a very low abundance of these sub-species present in the maternal gut and due to the nature of culturing they were not picked. If this subspecies was present in the offspring, re-culturing the maternal faecal sample in the presence of HMOs may be one way to select for this subspecies. One isolate, E002-T2.1, had SNP distance values >10,200 to the other isolates from participant E002 however had values <5 when compared to isolates from participant E005 suggesting clonal identity to these strains; this inconsistent result may have been down to human error at one stage during the preparation process for sequencing. No other individuals shared clonal strains between participants and they all had their own distinct *B. longum* strains; although in a recent publication it was noted that local populations have been shown to share strains between individuals however this paper used MAG-based (Metagenome-assembled genomes) sequencing rather than WGS, (158).

As seen in both **Figure 5.6/5.7** the isolates identified as *B. animalis* and the *B. animalis* subsp. *lactis* reference strains did not form any separate clades. All of the isolates, regardless of participant source or time-point, displayed a straight tree with the *lactis* strains indicating all isolates were *B. animalis* subsp. *lactis* rather than *B. animalis* subsp. *animalis* (confirmed with FastANI). The surprising lack of clades on the tree were supported with the data shown in **S.3** which revealed a large majority of the SNP distance values <40 including the reference strains used; the genomes revealed 1163 core genes were present. A large portion of these strains had SNP distance values <2 indicating many are clonal even though they were isolated from different patients at different timepoints. This suggests a common 'source' of strains that may be circulating between pregnant women however another possible explanation may be due to the recent emergence of this species causing there to be a lack of diversity between strains. Indeed, several of these strains had SNP distance values low enough to be considered clonal to 2 of the reference genomes, which are both probiotic strains. This included the most widely documented probiotic strain *B. animalis* subsp. *lactis* BB-12 which has been used in over 130 human clinical trials and is readily available in many probiotics which are commonly used (149). *B. animalis* was isolated from 10 out of the 12 participants and the participant information in **Table 3** was initially used to try and understand these data by comparing to the 2 participants (with no *B. animalis*) as controls such as diet, if the participants had a pet and previous probiotic use. However, no patterns were spotted which would explain the closely related isolates. Surprisingly, all 10 participants claimed to have not used any probiotics during

the study, and only 2 of those 10 had used them before. *B. animalis* subsp. *lactis* is present in many dairy products, (146), thus it is probable the high abundance of this subspecies in my dataset may be due to diet although the participants did confirm they had not consumed yogurt during the study. However, the close relation to the popular probiotic strains does imply the participants were ingesting these specific strains in some way. All of the participants were meat eaters, and another possibility could be the participants acquired these strains due to the use of probiotics in livestock, however as the meat would've been cooked this would be doubtful (159). Although unlikely, the participants did all attend the same hospital for check ins and to give birth and therefore, another possibility could be the participants all came in contact with this subspecies at the hospital. However, due to the need for sterile conditions in these environments and that *Bifidobacterium* is an anaerobe, I would assume this was not the case. Samples were also collected at trimester 1 from before the participants would have physically gone to the hospital for check-ups, suggesting this may not be a viable theory. Lee et al conducted a study looking at the genomes of different *Bifidobacterium* species and also found the four *B. animalis* subsp. *lactis* they looked at were very closely related with >99% sequence identity across all four genomes, (78), which suggests this sub-species group do not contain a large amount of genomic variability between them.

ARGs identified in the maternal gut microbiome could pose a threat for offspring due to HGT to potential pathogens, and also vertical transfer of MDR strains. Indeed, one study has already identified the ability to transfer tetracycline resistant bacterial strains, carrying the *tet(W)* gene, from mother to infant, (160-162), and our understanding of how these genes may populate the maternal microbiome is vital although tetracycline is also not widely used in humans causing the risk to health to be relatively low. Using the database ResFinder, information on the presence of 3 specific ARG: *aph(3')-Ia_7*, *erm*(X) and *tet(W)_4* was annotated onto the phylogenetic tree on **Figure 5.6** using blue bars. None of the isolates belonging to species *B. adolescentis, B. bifidum,* or *B. dentium* were found to have any of these genes encoded and only one *B. longum* isolate contained the one ARG *tet(W)_4*, which is the only E002 isolate that is clustered separately from the other from that participant. The *tet*(W) gene has been reported in both *B. bifidum* and *B. longum* previously with it seen at high frequency in *B. longum*, and therefore the lack of ARG found in both species was surprising (129), however this may be due to only three ARG were investigated. Aires et al also looked at different *tet* genes encoded in the genome of different *Bifidobacterium* species and also found none were encoded in their *B. adolescentis* and *B. dentium* isolates, although they did find a small amount

in their *B. bifidum* isolates and a high abundance in the *B. longum* isolates, (163). Every isolate identified as *B. animalis*, excluding isolate E030-T3-2, had the *tet*(W)_4 gene present in its genome. Another study conducted in 2021 yielded similar results with the *tet*(W) gene encoded in 41 out of 44 *B. animalis* subsp *lactis* strains however, further analysis revealed *B. animalis* subsp *lactis* encoded *tet*(W) is part of the ancient resistome and the risk of transfer is considered small, (164). Several studies have assessed the transferability of *tet*(W) in these strains, all of which were unsuccessful, (165, 166), which supports the data shown by Nohr-Meldgaard et al. Only one isolate, E009-T3-4, encoded for the ARG aph(3')-Ia_7, and *erm*(X) however this isolate was later discovered to have contamination and therefore it cannot be determined if this result is reliable. These studies, combined with my results suggests a low risk of transfer of these ARGs between strains, however it does point out that these species are very susceptible to antibiotics, demonstrating how these probiotic strains may be affected upon antibiotic administration.

6.2.4 Human Milk Oligosaccharides

Many species in the *Bifidobacterium* genus encode the functional capability to target and metabolise HMO's, contributing a major metabolic input to the infant host (130). Many of these species which contain the specialised gene clusters are often specific to the infant gut microbiome and tend to be absent in those associated with the adult gut, (167). However, considering the impact of the mother's microbiota on their offspring, there has been limited research on the presence of HMO clusters in maternally-derived *Bifidobacterium*. HMO's have been identified in the maternal circulation prior to birth and may offer a benefit for the mother during pregnancy, this study points at the potential for them to be also found in the gut during gestation which may affect the gut microbiota, (168). This would potentially be important – from the perspective of vertically 'seeding' the infant gut with strains that could utilise the infant (breast milk) diet. It has also been shown certain *Bifidobacterium* species employ the same enzymes for mucus degradation as it does for HMO utilisation, which may suggest another potential benefit for these clusters in the adult gut (169, 170).

Using an HMO database on Abricate-1.0.1, that was curated by Raymond Kiu, all 140 *Bifidobacterium* genomes were analysed for the presence of 6 known HMO clusters and a heatmap was created, using the ComplexHeatmap package version 2.10.0 (123), to visualise the presence of partially or fully present HMO clusters encoded in the genome, shown in

Figure 5.9. Further information on these clusters can be seen in *5.2.4 Human milk oligosaccharide clusters*.

The *B. longum* isolates encoded the highest number of HMO clusters with every isolate containing the full *nah* cluster and 8 of which, listed in *5.2.4 Human milk oligosaccharide clusters*, had all 6 clusters at least partially present in their genomes, these isolates are also clustered next to each other on the phylogenetic trees. These HMO clusters are used in the breakdown of lacto-N-tetraose (LNT) and lacto-N-neotetraose (LNnT), (James, 2016). *B. longum* subsp. *longum* is often found in the infant gut, (171), and therefore the high abundance of HMO clusters in the genomes of these isolates was expected. The *lac* cluster was only found partially present in 8 of these isolates and therefore may not be specific to this species.

B. bifidum is also associated with the infant gut microbiota and many studies have shown the ability of this species to break down HMO's, (169). All 4 *B. bifidum* isolates encoded the full *lac*, *nah,* and *lnp/glt* clusters and these isolates were the only ones to encode the full *lac* cluster out of all 140 isolates suggesting this cluster may be specific to *B. bifidum*. One study found *B. bifidum* differs in its pathway to breakdown certain HMOs from other *Bifidobacterium* species, their enzymes were found on the outside of the cell which allows for improved ability for cross feeding other microbiota members including different *Bifidobacterium* species, (James, 2016), this could explain the difference compared to the *B. longum* subsp. *longum* isolates.

The abundance of HMO clusters identified in the *B. adolescentis* isolates varied dependant on the participant the isolates were sourced from. Ten out of twelve of the isolates only fully encoded one HMO cluster with some also partially encoding up to 2 other clusters. Since *B. adolescentis* is more typically seen in the adult human gut compared to the infant gut, (146), this result was expected. Several other studies have shown *B. adolescentis* displays a limited capacity to break down HMO's, (172). The participant variability in this species group suggests HMO ability may also be strain specific rather than just species specific which has been shown in other *Bifidobacterium* species already, (173).

In contrast, to the *B. adolescentis* isolates, the 4 *B. dentium* strains had the same results even though they were isolated from 2 separate participants. None of these isolates contained any HMO clusters which were fully present, although the *lnt* and *nah* clusters, and the large *B.*

infantis cluster were partially present in their genomes. Although *B. dentium* has been isolated from infant faeces, many studies have shown the species lack of ability to utilise HMOs as a carbohydrate source, (174). Moya-Gonzalvez et al however found *B. dentium* was able to grow on LnT as an HMO source which would suggest why the *lnt* cluster was present in those isolates (175). These results may be due to *B. dentium* is more commonly found in the oral cavity.

None of the 57 *B. animalis* isolates fully encoded any of the HMO clusters and only 2 clusters, the *lac* and *lnt* clusters, were partially present in all of the isolates. The capability of the *B. animalis* species to metabolise HMOs has been previously described as poor, and similarly Lugli at el found only 15% of isolates contained a predicted lacto-*N*-biase-encoding gene in their genomes (172). One possible explanation for this loss in genomic ability when compared to other *Bifidobacterium* species may be due to the tendency of bacteria to undergo genomic reduction which may have occurred when *B. animalis* subsp. *lactis* evolved in the dairy fermentation process, (176). As these isolates are very genetically similar it is unsurprisingly the results were the same for each isolate.

Further investigation was conducted to try to understand any patterns that may have occurred related to the time point during gestation however, the data showed that the presence or absence of HMO clusters was determined by the species/strain and or the participant the isolates were sourced from. This conclusion is supported by many other studies that also found the presence of HMO's was strain specific, (177). Given more time functional studies would be completed using the isolates grown on different HMOs to identify if the encoded clusters are functional.

6.3.1 What would I do next?

This project was ultimately limited by COVID-19 and time constraints however, the data produced has laid an exciting foundation for further research, including more mechanistic studies.

Given more time, I would have liked to conduct further genomic based investigations into my isolates. This would have involved looking into the different glycoside-hydrolases that each species may encode (linking to abilities to digest a wider range of carbohydrates), and immune modulatory components such as exopolysaccharide capsules. This research would give me

more of an insight into why certain species may be more abundant in the maternal gut and what survival advantages they offer.

I would have liked to investigate the paired infant microbiomes from the mothers I had already studied. This would have allowed me to gain a greater understanding into species differences which are adult associated or infant associated, and how their genomes (and putative functionality) may differ. Using both FastANI and SNP distance values I could have also investigated how different species and strains are passed on from mother to offspring by identifying if any strains were clonal from both individuals and how what factors may impact this ability, such as birth mode and feeding method. The data used to study the presence of HMOs could also identify whether the species found in the infants are greater suited to break down HMOs compared to those already found in the maternal gut microbiota.

Once the metagenomic data for the samples I studied becomes available, I would have liked to compare this to my own genomic data to see how abundant the species I found were in the original samples. Culturomics, although incredibly important for understanding different species and strains can be limiting and there may have been other *Bifidobacterium* species that I was unable to isolate.

I had originally planned to perform several functional tests on the *Bifidobacterium* isolates in order to create in-depth profiles for the *Bif Bank*. These experiments would have investigated the isolate's reaction to acid shock, exposure to bile acids, and, oxygen exposure, as these traits are all key for development of next generation probiotics.

As mentioned in *1.2.2 The Maternal Microbiota*, previous research has highlighted that the maternal microbiome exhibits a shift in microbial composition as the pregnancy progresses, and this microbial shift may be driven by hormonal changes (44). During faecal culturing of my isolates, I was able to notice a distinct difference in the plates which were taken from the same participant at different time-points. Given more time, I would have liked to conduct a series of growth curve experiments with various hormones which are active during trimesters 2 and 3 on my isolates to see if the hormones have an impact on their growth.

To take this project further, it would be interesting to use these same techniques to study the maternal microbiome in different geographical locations and compare the abundance of *Bifidobacterium* species and strains across the world. This is important given differences in diet and other key microbiota-modulating factors and they reports differences in HMO composition between different maternal populations due to secretor status, (178).

6.3.2 Conclusion

The data from this project has provided genomic insights into the maternal microbiome throughout pregnancy and at birth. Five *Bifidobacterium* species were cultured and analysed from 159 isolates, with 88% of all *Bifidobacterium* isolates identified as *Bifidobacterium longum* subsp. *longum* or *Bifidobacterium animalis* subsp. *lactis*. The phylogenetic relationships between the species revealed very close genetic relatedness between several *B. animalis* subsp. *lactis* isolates, with many showing sequence similarity >99% with widely used probiotic strains despite participants not ingesting probiotic products. *B. longum* was the most abundant species identified, suggesting they offer a protective advantage during pregnancy, although the individual strains were person specific. The antimicrobial resistance gene *tet*(W)_4 was found to be encoded on all *B. animalis* isolates excluding one. Several HMO clusters were encoded in the genomes of the *Bifidobacterium* isolates with the most being identified in the *B. longum* and *B. bifidum* species and no HMO clusters were fully present in any of the *B. animalis* or *B. dentium* isolates.

Although colonisation and the composition of the infant gut has been widely documented, there is a distinct lack of research on the maternal gut microbiota despite studies suggesting it may impact pregnancy outcomes, and short/ long term health consequences for the maternal offspring. This study will contribute to our further understanding of the healthy maternal gut microbiota and how it's *Bifidobacterium* composition may alter over the gestation period. Our knowledge on the important topic may aid the development of improved maternal health and the potential development of pregnancy focused probiotics.

7.0 Acknowledgements

I would like to take the time to mention the people that made writing this thesis possible. I have thoroughly enjoyed my time spent working on this project and this amazing experience has all been down to the masses of support from my family and peers.

First and foremost, I would like to thank Lindsay, you have not only been an incredible supervisor these past 18 months, but you have also been a huge inspiration for me. I feel especially lucky to have had the pleasure of working with you on this project and as a young woman in STEM I feel grateful to have such an amazing role model to look up to.

Secondly, I would like to thank Nancy who couldn't have been a better mentor if she tried. Thank you so much for taking the time out of your crazy schedule to teach me and coach me every step of the way. I couldn't ask for a more patient and selfless individual who went above and beyond to help me throughout this project; to work with two such courageous women has been an honour.

I would also like to thank the rest of the Hall lab group whose support and great ideas has been vital for those many tricky questions I came to face. With a big thank you to both Raymond and Matthew who were a huge help with the bioinformatics side of my project whether that was always being on hand when my code didn't work or aiding me in creating some of the figures in this thesis; and *Bifidobacterium* genome expert Magda who was always able to answer my frequent questions. With that in hand I would like to extend my gratitude to the members of the PEARL team, especially both Sarah and Tom, who continuously hunt down samples for me or provided me with the many answers I had on our participants and, the PEARL participants who provided those very samples this project was based on.

My masters project would not have been such a wonderful experience if it wasn't for the very special Robinson lab members both past and present. Thank you to Stephen for the constant emotional support this past year, I have felt especially lucky to have a supervisor also part of the LGBTQ+ community and I cannot thank you enough for the warm welcome into your group. I would also like to add a special mention to Chris Benwell for taking on the role as the lab big brother – our daily chats have often been what I looked forward to most, both Luke and Alicia for providing constant entertainment and chaos and the rest of the lab group for

providing continuous support in and out of work; the many caffeination breaks have been both welcomed and often needed.

Lastly, I would like to thank my family and friends for always being there this past year. Thank you to both my parents for always pushing me to acquire everything I had ever dreamed of; I would not be here if it wasn't for both of your undying support day in and day out. Thank you to my eldest brother Michael, who was always on hand to video call me whenever I needed help coding or for a quick pep talk with my little nephew, and to my middle brother Lee, who has encouraged me every step of the way throughout my career into science.

This masters project has been both a pleasure to write and complete and I am incredibly grateful for my time at UEA.

8.0 References

1. Power SE, O'Toole PW, Stanton C, Ross RP, Fitzgerald GF. Intestinal microbiota, diet and health. Br J Nutr. 2014;111(3):387-402.
2. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464(7285):59-65.
3. Baothman OA, Zamzami MA, Taher I, Abubaker J, Abu-Farha M. The role of Gut Microbiota in the development of obesity and Diabetes. Lipids Health Dis. 2016;15:108.
4. Kaser A, Zeissig S, Blumberg RS. Inflammatory bowel disease. Annu Rev Immunol. 2010;28:573-621.
5. Meng C, Bai C, Brown TD, Hood LE, Tian Q. Human Gut Microbiota and Gastrointestinal Cancer. Genomics Proteomics Bioinformatics. 2018;16(1):33-49.
6. Jiang C, Li G, Huang P, Liu Z, Zhao B. The Gut Microbiota and Alzheimer's Disease. J Alzheimers Dis. 2017;58(1):1-15.
7. Morkl S, Butler MI, Holl A, Cryan JF, Dinan TG. Probiotics and the Microbiota-Gut-Brain Axis: Focus on Psychiatry. Curr Nutr Rep. 2020;9(3):171-82.
8. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the human infant intestinal microbiota. PLoS Biol. 2007;5(7):e177.
9. Hasan N, Yang H. Factors affecting the composition of the gut microbiota, and its modulation. PeerJ. 2019;7:e7502.
10. Molina-Torres G, Rodriguez-Arrastia M, Roman P, Sanchez-Labraca N, Cardona D. Stress and the gut microbiota-brain axis. Behav Pharmacol. 2019;30(2 and 3-Spec Issue):187-200.
11. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Nageshwar Reddy D. Role of the normal gut microbiota. World J Gastroenterol. 2015;21(29):8787-803.
12. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. Science. 2008;320(5883):1647-51.
13. Bergstrom A, Skov TH, Bahl MI, Roager HM, Christensen LB, Ejlerskov KT, et al. Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. Appl Environ Microbiol. 2014;80(9):2889-900.
14. Arrieta MC, Stiemsma LT, Amenyogbe N, Brown EM, Finlay B. The intestinal microbiome in early life: health and disease. Front Immunol. 2014;5:427.

15. Eisenhauer N, Scheu S, Jousset A. Bacterial diversity stabilizes community productivity. PLoS One. 2012;7(3):e34517.
16. Tetz VV. The pangenome concept: a unifying view of genetic information. Med Sci Monit. 2005;11(7):HY24-9.
17. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A. 2005;102(39):13950-5.
18. Brockhurst MA, Harrison E, Hall JPJ, Richards T, McNally A, MacLean C. The Ecology and Evolution of Pangenomes. Curr Biol. 2019;29(20):R1094-R103.
19. Soucy SM, Huang J, Gogarten JP. Horizontal gene transfer: building the web of life. Nat Rev Genet. 2015;16(8):472-82.
20. Choudhuri S. Bioinformatics for beginners : genes, genomes, molecular evolution, databases and analytical tools. Amsterdam: Academic Press; 2014.
21. Oliphant K, Allen-Vercoe E. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. Microbiome. 2019;7(1):91.
22. Hill MJ. Intestinal flora and endogenous vitamin synthesis. Eur J Cancer Prev. 1997;6 Suppl 1:S43-5.
23. Dalile B, Van Oudenhove L, Vervliet B, Verbeke K. The role of short-chain fatty acids in microbiota-gut-brain communication. Nat Rev Gastroenterol Hepatol. 2019;16(8):461-78.
24. Machiels K, Joossens M, Sabino J, De Preter V, Arijs I, Eeckhaut V, et al. A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. Gut. 2014;63(8):1275-83.
25. Sampson TR, Mazmanian SK. Control of brain development, function, and behavior by the microbiome. Cell Host Microbe. 2015;17(5):565-76.
26. Umesaki Y, Okada Y, Matsumoto S, Imaoka A, Setoyama H. Segmented filamentous bacteria are indigenous intestinal bacteria that activate intraepithelial lymphocytes and induce MHC class II molecules and fucosyl asialo GM1 glycolipids on the small intestinal epithelial cells in the ex-germ-free mouse. Microbiol Immunol. 1995;39(8):555-62.
27. Cebra JJ. Influences of microbiota on intestinal immune system development. Am J Clin Nutr. 1999;69(5):1046S-51S.
28. Chang PV, Hao L, Offermanns S, Medzhitov R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. Proc Natl Acad Sci U S A. 2014;111(6):2247-52.

29. Thavagnanam S, Fleming J, Bromley A, Shields MD, Cardwell CR. A meta-analysis of the association between Caesarean section and childhood asthma. Clin Exp Allergy. 2008;38(4):629-33.
30. Cardwell CR, Stene LC, Joner G, Cinek O, Svensson J, Goldacre MJ, et al. Caesarean section is associated with an increased risk of childhood-onset type 1 diabetes mellitus: a meta-analysis of observational studies. Diabetologia. 2008;51(5):726-35.
31. Buffie CG, Pamer EG. Microbiota-mediated colonization resistance against intestinal pathogens. Nat Rev Immunol. 2013;13(11):790-801.
32. Santagati M, Scillato M, Patane F, Aiello C, Stefani S. Bacteriocin-producing oral streptococci and inhibition of respiratory pathogens. FEMS Immunol Med Microbiol. 2012;65(1):23-31.
33. Slizewska K, Markowiak-Kopec P, Slizewska W. The Role of Probiotics in Cancer Prevention. Cancers (Basel). 2020;13(1).
34. Schrezenmeir J, de Vrese M. Probiotics, prebiotics, and synbiotics--approaching a definition. Am J Clin Nutr. 2001;73(2 Suppl):361S-4S.
35. Hennet T, Borsig L. Breastfed at Tiffany's. Trends Biochem Sci. 2016;41(6):508-18.
36. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, et al. Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol. 2014;11(8):506-14.
37. Yang B, Lu P, Li MX, Cai XL, Xiong WY, Hou HJ, et al. A meta-analysis of the effects of probiotics and synbiotics in children with acute diarrhea. Medicine (Baltimore). 2019;98(37):e16618.
38. Brantsaeter AL, Myhre R, Haugen M, Myking S, Sengpiel V, Magnus P, et al. Intake of probiotic food and risk of preeclampsia in primiparous women: the Norwegian Mother and Child Cohort Study. Am J Epidemiol. 2011;174(7):807-15.
39. Cordaillat-Simmons M, Rouanet A, Pot B. Live biotherapeutic products: the importance of a defined regulatory framework. Exp Mol Med. 2020;52(9):1397-406.
40. Quraishi MN, Widlak M, Bhala N, Moore D, Price M, Sharma N, et al. Systematic review with meta-analysis: the efficacy of faecal microbiota transplantation for the treatment of recurrent and refractory Clostridium difficile infection. Aliment Pharmacol Ther. 2017;46(5):479-93.

41. Costello SP, Soo W, Bryant RV, Jairath V, Hart AL, Andrews JM. Systematic review with meta-analysis: faecal microbiota transplantation for the induction of remission for active ulcerative colitis. Aliment Pharmacol Ther. 2017;46(3):213-24.
42. Lockitch G. Clinical biochemistry of pregnancy. Crit Rev Clin Lab Sci. 1997;34(1):67-139.
43. Soma-Pillay P, Nelson-Piercy C, Tolppanen H, Mebazaa A. Physiological changes in pregnancy. Cardiovasc J Afr. 2016;27(2):89-94.
44. Koren O, Goodrich JK, Cullender TC, Spor A, Laitinen K, Backhed HK, et al. Host remodeling of the gut microbiome and metabolic changes during pregnancy. Cell. 2012;150(3):470-80.
45. Nuriel-Ohayon M, Neuman H, Koren O. Microbial Changes during Pregnancy, Birth, and Infancy. Front Microbiol. 2016;7:1031.
46. Collado MC, Isolauri E, Laitinen K, Salminen S. Distinct composition of gut microbiota during pregnancy in overweight and normal-weight women. Am J Clin Nutr. 2008;88(4):894-9.
47. Zietek M, Celewicz Z, Szczuko M. Short-Chain Fatty Acids, Maternal Microbiota and Metabolism in Pregnancy. Nutrients. 2021;13(4).
48. Mulak A, Tache Y, Larauche M. Sex hormones in the modulation of irritable bowel syndrome. World J Gastroenterol. 2014;20(10):2433-48.
49. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A. 2010;107(26):11971-5.
50. O'Sullivan A, Farver M, Smilowitz JT. The Influence of Early Infant-Feeding Practices on the Intestinal Microbiome and Body Composition in Infants. Nutr Metab Insights. 2015;8(Suppl 1):1-9.
51. Rouge C, Goldenberg O, Ferraris L, Berger B, Rochat F, Legrand A, et al. Investigation of the intestinal microbiota in preterm infants using different methods. Anaerobe. 2010;16(4):362-70.
52. Houttu N, Mokkala K, Laitinen K. Overweight and obesity status in pregnant women are related to intestinal microbiota and serum metabolic and inflammatory profiles. Clin Nutr. 2018;37(6 Pt A):1955-66.
53. Santacruz A, Collado MC, Garcia-Valdes L, Segura MT, Martin-Lagos JA, Anjos T, et al. Gut microbiota composition is associated with body weight, weight gain and biochemical parameters in pregnant women. Br J Nutr. 2010;104(1):83-92.

54. Cani PD, Neyrinck AM, Fava F, Knauf C, Burcelin RG, Tuohy KM, et al. Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. Diabetologia. 2007;50(11):2374-83.
55. Gomez-Arango LF, Barrett HL, McIntyre HD, Callaway LK, Morrison M, Dekker Nitert M, et al. Increased Systolic and Diastolic Blood Pressure Is Associated With Altered Gut Microbiota Composition and Butyrate Production in Early Pregnancy. Hypertension. 2016;68(4):974-81.
56. Rana S, Lemoine E, Granger JP, Karumanchi SA. Preeclampsia: Pathophysiology, Challenges, and Perspectives. Circ Res. 2019;124(7):1094-112.
57. Lv LJ, Li SH, Li SC, Zhong ZC, Duan HL, Tian C, et al. Early-Onset Preeclampsia Is Associated With Gut Microbial Alterations in Antepartum and Postpartum Women. Front Cell Infect Microbiol. 2019;9:224.
58. Chen H, Li J, Li N, Liu H, Tang J. Increased circulating trimethylamine N-oxide plays a contributory role in the development of endothelial dysfunction and hypertension in the RUPP rat model of preeclampsia. Hypertens Pregnancy. 2019;38(2):96-104.
59. Hu M, Eviston D, Hsu P, Marino E, Chidgey A, Santner-Nanan B, et al. Decreased maternal serum acetate and impaired fetal thymic and regulatory T cell development in preeclampsia. Nat Commun. 2019;10(1):3031.
60. Di Simone N, Santamaria Ortiz A, Specchia M, Tersigni C, Villa P, Gasbarrini A, et al. Recent Insights on the Maternal Microbiota: Impact on Pregnancy Outcomes. Front Immunol. 2020;11:528202.
61. Collado MC, Isolauri E, Laitinen K, Salminen S. Effect of mother's weight on infant's microbiota acquisition, composition, and activity during early infancy: a prospective follow-up study initiated in early pregnancy. Am J Clin Nutr. 2010;92(5):1023-30.
62. Bjorksten B, Sepp E, Julge K, Voor T, Mikelsaar M. Allergy development and the intestinal microflora during the first year of life. J Allergy Clin Immunol. 2001;108(4):516-20.
63. Collado MC, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y. Specific duodenal and faecal bacterial groups associated with paediatric coeliac disease. J Clin Pathol. 2009;62(3):264-9.
64. Bookstaver PB, Bland CM, Griffin B, Stover KR, Eiland LS, McLaughlin M. A Review of Antibiotic Use in Pregnancy. Pharmacotherapy. 2015;35(11):1052-62.
65. Yoon MY, Yoon SS. Disruption of the Gut Ecosystem by Antibiotics. Yonsei Med J. 2018;59(1):4-12.

66. Metsala J, Lundqvist A, Virta LJ, Kaila M, Gissler M, Virtanen SM. Prenatal and post-natal exposure to antibiotics and risk of asthma in childhood. Clin Exp Allergy. 2015;45(1):137-45.

67. Mezzelani A, Landini M, Facchiano F, Raggi ME, Villa L, Molteni M, et al. Environment, dysbiosis, immunity and sex-specific susceptibility: a translational hypothesis for regressive autism pathogenesis. Nutr Neurosci. 2015;18(4):145-61.

68. Ferretti P, Pasolli E, Tett A, Asnicar F, Gorfer V, Fedi S, et al. Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. Cell Host Microbe. 2018;24(1):133-45 e5.

69. Pessa-Morikawa T, Husso A, Karkkainen O, Koistinen V, Hanhineva K, Iivanainen A, et al. Maternal microbiota-derived metabolic profile in fetal murine intestine, brain and placenta. BMC Microbiol. 2022;22(1):46.

70. Browne HP, Shao Y, Lawley TD. Mother-infant transmission of human microbiota. Curr Opin Microbiol. 2022;69:102173.

71. Shao Y, Forster SC, Tsaliki E, Vervier K, Strang A, Simpson N, et al. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. Nature. 2019;574(7776):117-21.

72. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. Nat Med. 2016;22(7):713-22.

73. Amenyogbe N, Kollmann TR, Ben-Othman R. Early-Life Host-Microbiome Interphase: The Key Frontier for Immune Development. Front Pediatr. 2017;5:111.

74. Korpela K, Helve O, Kolho KL, Saisto T, Skogberg K, Dikareva E, et al. Maternal Fecal Microbiota Transplantation in Cesarean-Born Infants Rapidly Restores Normal Gut Microbial Development: A Proof-of-Concept Study. Cell. 2020;183(2):324-34 e5.

75. Wilson BC, Butler EM, Grigg CP, Derraik JGB, Chiavaroli V, Walker N, et al. Oral administration of maternal vaginal microbes at birth to restore gut microbiome development in infants born by caesarean section: A pilot randomised placebo-controlled trial. EBioMedicine. 2021;69:103443.

76. Turroni F, van Sinderen D, Ventura M. Genomics and ecological overview of the genus Bifidobacterium. Int J Food Microbiol. 2011;149(1):37-44.

77. Miyake T, Watanabe K, Watanabe T, Oyaizu H. Phylogenetic analysis of the genus Bifidobacterium and related genera based on 16S rDNA sequences. Microbiol Immunol. 1998;42(10):661-7.

78. Lee JH, O'Sullivan DJ. Genomic insights into bifidobacteria. Microbiol Mol Biol Rev. 2010;74(3):378-416.

79. O'Callaghan A, van Sinderen D. Bifidobacteria and Their Role as Members of the Human Gut Microbiota. Front Microbiol. 2016;7:925.

80. Sela DA, Mills DA. Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides. Trends Microbiol. 2010;18(7):298-307.

81. Slavin J. Fiber and prebiotics: mechanisms and health benefits. Nutrients. 2013;5(4):1417-35.

82. Shokryazdan P, Faseleh Jahromi M, Navidshad B, Liang JB. Effects of prebiotics on immune system and cytokine expression. Med Microbiol Immunol. 2017;206(1):1-9.

83. Wong JM, de Souza R, Kendall CW, Emam A, Jenkins DJ. Colonic health: fermentation and short chain fatty acids. J Clin Gastroenterol. 2006;40(3):235-43.

84. Noda H, Akasaka N, Ohsugi M. Biotin production by bifidobacteria. J Nutr Sci Vitaminol (Tokyo). 1994;40(2):181-8.

85. Yoshii K, Hosomi K, Sawane K, Kunisawa J. Metabolism of Dietary and Microbial Vitamin B Family in the Regulation of Host Immunity. Front Nutr. 2019;6:48.

86. Solopova A, Bottacini F, Venturi Degli Esposti E, Amaretti A, Raimondi S, Rossi M, et al. Riboflavin Biosynthesis and Overproduction by a Derivative of the Human Gut Commensal Bifidobacterium longum subsp. infantis ATCC 15697. Front Microbiol. 2020;11:573335.

87. Yang SC, Lin CH, Sung CT, Fang JY. Antibacterial activities of bacteriocins: application in foods and pharmaceuticals. Front Microbiol. 2014;5:241.

88. Darvishi N, Fard NA, Sadrnia M. Genomic and proteomic comparisons of bacteriocins in probiotic species Lactobacillus and Bifidobacterium and inhibitory ability of Escherichia coli MG 1655. Biotechnol Rep (Amst). 2021;31:e00654.

89. Javvadi SG, Kujawska M, Papp D, Gontarczyk AM, Jordan A, Lawson MAE, et al. A novel bacteriocin produced by Bifidobacterium longum subsp. infantis has dual antimicrobial and immunomodulatory activity. bioRxiv. 2022:2022.01.27.477972.

90. Klijn A, Mercenier A, Arigoni F. Lessons from the genomes of bifidobacteria. FEMS Microbiology Reviews. 2005;29(3):491-509.

91. Kiu R, Treveil A, Harnisch LC, Caim S, Leclaire C, van Sinderen D, et al. Bifidobacterium breve UCC2003 Induces a Distinct Global Transcriptomic Program in Neonatal Murine Intestinal Epithelial Cells. iScience. 2020;23(7):101336.

92. Ratsika A, Codagnone MC, O'Mahony S, Stanton C, Cryan JF. Priming for Life: Early Life Nutrition and the Microbiota-Gut-Brain Axis. Nutrients. 2021;13(2).
93. Palma GD, Capilla A, Nova E, Castillejo G, Varea V, Pozo T, et al. Influence of milk-feeding type and genetic risk of developing coeliac disease on intestinal microbiota of infants: the PROFICEL study. PLoS One. 2012;7(2):e30791.
94. Ly NP, Litonjua A, Gold DR, Celedon JC. Gut microbiota, probiotics, and vitamin D: interrelated exposures influencing allergy, asthma, and obesity? J Allergy Clin Immunol. 2011;127(5):1087-94; quiz 95-6.
95. Amarasekara R, Jayasekara RW, Senanayake H, Dissanayake VH. Microbiome of the placenta in pre-eclampsia supports the role of bacteria in the multifactorial cause of pre-eclampsia. J Obstet Gynaecol Res. 2015;41(5):662-9.
96. Milani C, Duranti S, Bottacini F, Casey E, Turroni F, Mahony J, et al. The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. Microbiol Mol Biol Rev. 2017;81(4).
97. Lopez-Tello J, Schofield Z, Kiu R, Dalby MJ, van Sinderen D, Gall GL, et al. Maternal microbiota Bifidobacterium promotes placental morphogenesis, nutrient transport and fetal growth in mice. bioRxiv. 2022:2021.07.23.453407.
98. Chen J, Chen X, Ho CL. Recent Development of Probiotic Bifidobacteria for Treating Human Diseases. Front Bioeng Biotechnol. 2021;9:770248.
99. Yanokura E, Oki K, Makino H, Modesto M, Pot B, Mattarelli P, et al. Subspeciation of Bifidobacterium longum by multilocus approaches and amplified fragment length polymorphism: Description of B. longum subsp. suillum subsp. nov., isolated from the faeces of piglets. Syst Appl Microbiol. 2015;38(5):305-14.
100. Chichlowski M, Shah N, Wampler JL, Wu SS, Vanderhoof JA. Bifidobacterium longum Subspecies infantis (B. infantis) in Pediatric Nutrition: Current State of Knowledge. Nutrients. 2020;12(6).
101. Chichlowski M, German JB, Lebrilla CB, Mills DA. The influence of milk oligosaccharides on microbiota of infants: opportunities for formulas. Annu Rev Food Sci Technol. 2011;2:331-51.
102. Qin F, Du S, Zhang Z, Ying H, Wu Y, Zhao G, et al. Newly identified HMO-2011-type phages reveal genomic diversity and biogeographic distributions of this marine viral group. ISME J. 2022;16(5):1363-75.

103. German JB, Freeman SL, Lebrilla CB, Mills DA. Human milk oligosaccharides: evolution, structures and bioselectivity as substrates for intestinal bacteria. Nestle Nutr Workshop Ser Pediatr Program. 2008;62:205-18; discussion 18-22.
104. Marcobal A, Sonnenburg JL. Human milk oligosaccharide consumption by intestinal microbiota. Clin Microbiol Infect. 2012;18 Suppl 4:12-5.
105. Wicinski M, Sawicka E, Gebalski J, Kubiak K, Malinowski B. Human Milk Oligosaccharides: Health Benefits, Potential Applications in Infant Formulas, and Pharmacology. Nutrients. 2020;12(1).
106. Morrow AL, Ruiz-Palacios GM, Altaye M, Jiang X, Guerrero ML, Meinzen-Derr JK, et al. Human milk oligosaccharide blood group epitopes and innate immune protection against campylobacter and calicivirus diarrhea in breastfed infants. Adv Exp Med Biol. 2004;554:443-6.
107. Hirschmugl B, Brandl W, Csapo B, van Poppel M, Kofeler H, Desoye G, et al. Evidence of Human Milk Oligosaccharides in Cord Blood and Maternal-to-Fetal Transport across the Placenta. Nutrients. 2019;11(11).
108. Zivkovic AM, Barile D. Bovine milk as a source of functional oligosaccharides for improving human health. Adv Nutr. 2011;2(3):284-9.
109. Nuriel-Ohayon M, Neuman H, Ziv O, Belogolovski A, Barsheshet Y, Bloch N, et al. Progesterone Increases Bifidobacterium Relative Abundance during Late Pregnancy. Cell Rep. 2019;27(3):730-6 e3.
110. Garcia Trallero O, Herrera Serrano L, Bibian Ingles M, Roche Valles D, Rodriguez AM. Effect of the administration of a probiotic with a combination of Lactobacillus and Bifidobacterium strains on antibiotic-associated diarrhea. Rev Esp Quimioter. 2019;32(3):268-72.
111. Martin-Pelaez S, Cano-Ibanez N, Pinto-Gallardo M, Amezcua-Prieto C. The Impact of Probiotics, Prebiotics, and Synbiotics during Pregnancy or Lactation on the Intestinal Microbiota of Children Born by Cesarean Section: A Systematic Review. Nutrients. 2022;14(2).
112. Alcon-Giner C, Dalby MJ, Caim S, Ketskemety J, Shaw A, Sim K, et al. Microbiota Supplementation with Bifidobacterium and Lactobacillus Modifies the Preterm Infant Gut Microbiota and Metabolome: An Observational Study. Cell Rep Med. 2020;1(5):100077.
113. Phillips S, Watt R, Atkinson T, Savva GM, Hayhoe A, Hall LJ, et al. The Pregnancy and EARly Life study (PEARL) - a longitudinal study to understand how gut microbes

contribute to maintaining health during pregnancy and early life. BMC Pediatr. 2021;21(1):357.

114. Baker DJ, Aydin A, Le-Viet T, Kay GL, Rudder S, de Oliveira Martins L, et al. CoronaHiT: high-throughput sequencing of SARS-CoV-2 genomes. Genome Med. 2021;13(1):21.

115. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90.

116. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. Curr Protoc Bioinformatics. 2020;70(1):e102.

117. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019;36(6):1925-7.

118. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.

119. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.

120. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb Genom. 2016;2(4):e000056.

121. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37(5):1530-4.

122. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293-W6.

123. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847-9.

124. Batt CA, Tortorello ML. Encyclopedia of food microbiology. Second edition ed. Amsterdam: Elsevier/Academic Press; 2014. 3 volumes : illustrations (black and white, and colour), 1 map (colour) p.

125. Milani C, Lugli GA, Duranti S, Turroni F, Bottacini F, Mangifesta M, et al. Genomic encyclopedia of type strains of the genus Bifidobacterium. Appl Environ Microbiol. 2014;80(20):6290-302.

126. T S. Abricate [Available from: https://github.com/tseeman/abricate/.

127. Lu W, Li K, Huang J, Sun Z, Li A, Liu H, et al. Identification and characteristics of a novel aminoglycoside phosphotransferase, APH(3')-IId, from an MDR clinical isolate of Brucella intermedia. J Antimicrob Chemother. 2021;76(11):2787-94.
128. Hays C, Lienhard R, Auzou M, Barraud O, Guerin F, Ploy MC, et al. Erm(X)-mediated resistance to macrolides, lincosamides and streptogramins in Actinobaculum schaalii. J Antimicrob Chemother. 2014;69(8):2056-60.
129. Ammor MS, Florez AB, Alvarez-Martin P, Margolles A, Mayo B. Analysis of tetracycline resistance tet(W) genes and their flanking sequences in intestinal Bifidobacterium species. J Antimicrob Chemother. 2008;62(4):688-93.
130. Kijner S, Cher A, Yassour M. The Infant Gut Commensal Bacteroides dorei Presents a Generalized Transcriptional Response to Various Human Milk Oligosaccharides. Front Cell Infect Microbiol. 2022;12:854122.
131. James K, Motherway MO, Bottacini F, van Sinderen D. Bifidobacterium breve UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through overlapping, yet distinct pathways. Sci Rep. 2016;6:38560.
132. Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, et al. The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. Proc Natl Acad Sci U S A. 2008;105(48):18964-9.
133. Bonet M, Ota E, Chibueze CE, Oladapo OT. Routine antibiotic prophylaxis after normal vaginal birth for reducing maternal infectious morbidity. Cochrane Database Syst Rev. 2017;11(11):CD012137.
134. Chen CM, Chou HC, Yang YSH. Maternal Antibiotic Treatment Disrupts the Intestinal Microbiota and Intestinal Development in Neonatal Mice. Front Microbiol. 2021;12:684233.
135. Dunn AB, Jordan S, Baker BJ, Carlson NS. The Maternal Infant Microbiome: Considerations for Labor and Birth. MCN Am J Matern Child Nurs. 2017;42(6):318-25.
136. Caballero B, Finglas PM, Toldrá F, ScienceDirect, Elsevier. Encyclopedia of food and health. Oxford, England
London: Academic Press, an imprint of Elsevier Ltd
ScienceDirect; 2016.
137. Lawson MAE, O'Neill IJ, Kujawska M, Gowrinadh Javvadi S, Wijeyesekera A, Flegg Z, et al. Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. ISME J. 2020;14(2):635-48.

138. Mattarelli P, Biavati B, Holzapfel WH, Wood BJB, ScienceDirect. The bifidobacteria and related organisms : biology, taxonomy, applications. London, United Kingdom: Academic Press is an imprint of Elsevier; 2017.
139. Gueimonde M, Sanchez B, C GdLR-G, Margolles A. Antibiotic resistance in probiotic bacteria. Front Microbiol. 2013;4:202.
140. Hu X, Lin B, Luo M, Zheng X, Zhang H. The isolation, identification, physiological property of pig-isolate Clostridium butyricum LY33 using lactic acid and its effects on intestinal function of weaned piglets. Italian Journal of Animal Science. 2019;18(1):910-21.
141. Kutsuna R, Miyoshi-Akiyama T, Mori K, Hayashi M, Tomida J, Morita Y, et al. Description of Paraclostridium bifermentans subsp. muricolitidis subsp. nov., emended description of Paraclostridium bifermentans (Sasi Jyothsna et al., 2016), and creation of Paraclostridium bifermentans subsp. bifermentans subsp. nov. Microbiol Immunol. 2019;63(1):1-10.
142. Mehdizadeh Gohari I, M AN, Li J, Shrestha A, Uzal F, B AM. Pathogenicity and virulence of Clostridium perfringens. Virulence. 2021;12(1):723-53.
143. Zhao H, Wang J, Peng Y, Cai X, Liu Y, Huang W, et al. Genomic insights from Paraclostridium bifermentans HD0315_2: General features and pathogenic potential. Front Microbiol. 2022;13:928153.
144. Shaw AG, Cornwell E, Sim K, Thrower H, Scott H, Brown JCS, et al. Dynamics of toxigenic Clostridium perfringens colonisation in a cohort of prematurely born neonatal infants. BMC Pediatr. 2020;20(1):75.
145. Kiu R, Caim S, Painset A, Pickard D, Swift C, Dougan G, et al. Phylogenomic analysis of gastroenteritis-associated Clostridium perfringens in England and Wales over a 7-year period indicates distribution of clonal toxigenic strains in multiple outbreaks and extensive involvement of enterotoxin-encoding (CPE) plasmids. Microb Genom. 2019;5(10).
146. Derrien M, Turroni F, Ventura M, van Sinderen D. Insights into endogenous Bifidobacterium species in the human gut microbiota during adulthood. Trends Microbiol. 2022;30(10):940-7.
147. Arboleya S, Bottacini F, O'Connell-Motherway M, Ryan CA, Ross RP, van Sinderen D, et al. Gene-trait matching across the Bifidobacterium longum pan-genome reveals considerable diversity in carbohydrate catabolism among human infant strains. BMC Genomics. 2018;19(1):33.
148. Cheng J, Laitila A, Ouwehand AC. Bifidobacterium animalis subsp. lactis HN019 Effects on Gut Health: A Review. Front Nutr. 2021;8:790561.

149. Jungersen M, Wind A, Johansen E, Christensen JE, Stuer-Lauridsen B, Eskesen D. The Science behind the Probiotic Strain Bifidobacterium animalis subsp. lactis BB-12((R)). Microorganisms. 2014;2(2):92-110.
150. Yang B, Ding M, Chen Y, Han F, Yang C, Zhao J, et al. Development of gut microbiota and bifidobacterial communities of neonates in the first 6 weeks and their inheritance from mother. Gut Microbes. 2021;13(1):1-13.
151. Junick J, Blaut M. Quantification of human fecal bifidobacterium species by use of quantitative real-time PCR analysis targeting the groEL gene. Appl Environ Microbiol. 2012;78(8):2613-22.
152. Oliver A, Chase AB, Weihe C, Orchanian SB, Riedel SF, Hendrickson CL, et al. High-Fiber, Whole-Food Dietary Intervention Alters the Human Gut Microbiome but Not Fecal Short-Chain Fatty Acids. mSystems. 2021;6(2).
153. Long X, Liu D, Gao Q, Ni J, Qian L, Ni Y, et al. Bifidobacterium adolescentis Alleviates Liver Steatosis and Steatohepatitis by Increasing Fibroblast Growth Factor 21 Sensitivity. Front Endocrinol (Lausanne). 2021;12:773340.
154. Fernandez-Duarte KP, Olaya-Galan NN, Salas-Cardenas SP, Lopez-Rozo J, Gutierrez-Fernandez MF. Bifidobacterium adolescentis (DSM 20083) and Lactobacillus casei (Lafti L26-DSL): Probiotics Able to Block the In Vitro Adherence of Rotavirus in MA104 Cells. Probiotics Antimicrob Proteins. 2018;10(1):56-63.
155. Choi JH, Lee KM, Lee MK, Cha CJ, Kim GB. Bifidobacterium faecale sp. nov., isolated from human faeces. Int J Syst Evol Microbiol. 2014;64(Pt 9):3134-9.
156. Engevik MA, Danhof HA, Hall A, Engevik KA, Horvath TD, Haidacher SJ, et al. The metabolic profile of Bifidobacterium dentium reflects its status as a human gut commensal. BMC Microbiol. 2021;21(1):154.
157. Yao S, Zhao Z, Wang W, Liu X. Bifidobacterium Longum: Protection against Inflammatory Bowel Disease. J Immunol Res. 2021;2021:8030297.
158. Valles-Colomer M, Blanco-Miguez A, Manghi P, Asnicar F, Dubois L, Golzato D, et al. The person-to-person transmission landscape of the gut and oral microbiomes. Nature. 2023;614(7946):125-35.
159. Mahesh MS, Mohanta RK, Patra AK. Probiotics in Livestock and Poultry Nutrition and Health. In: Goel G, Kumar A, editors. Advances in Probiotics for Sustainable Food and Medicine. Singapore: Springer Singapore; 2021. p. 149-79.
160. Alicea-Serrano AM, Contreras M, Magris M, Hidalgo G, Dominguez-Bello MG. Tetracycline resistance genes acquired at birth. Arch Microbiol. 2013;195(6):447-51.

161. Kozak K, Charbonneau D, Sanozky-Dawes R, Klaenhammer T. Characterization of bacterial isolates from the microbiota of mothers' breast milk and their infants. Gut Microbes. 2015;6(6):341-51.
162. Das L, Virmani R, Sharma V, Rawat D, Singh Y. Human Milk Microbiota: Transferring the Antibiotic Resistome to Infants. Indian J Microbiol. 2019;59(4):410-6.
163. Aires J, Doucet-Populaire F, Butel MJ. Tetracycline resistance mediated by tet(W), tet(M), and tet(O) genes of Bifidobacterium isolates from humans. Appl Environ Microbiol. 2007;73(8):2751-4.
164. Nohr-Meldgaard K, Struve C, Ingmer H, Agerso Y. The Tetracycline Resistance Gene, tet(W) in Bifidobacterium animalis subsp. lactis Follows Phylogeny and Differs From tet(W) in Other Species. Front Microbiol. 2021;12:658943.
165. Gueimonde M, Florez AB, van Hoek AH, Stuer-Lauridsen B, Stroman P, de los Reyes-Gavilan CG, et al. Genetic basis of tetracycline resistance in Bifidobacterium animalis subsp. lactis. Appl Environ Microbiol. 2010;76(10):3364-9.
166. Polit A, Yang H, Amund D. Investigating the transmissibility of tet(W) in bifidobacteria exposed to acid and bile stress. Biosci Microbiota Food Health. 2018;37(2):39-43.
167. Turroni F, Milani C, Duranti S, Mahony J, van Sinderen D, Ventura M. Glycan Utilization and Cross-Feeding Activities by Bifidobacteria. Trends Microbiol. 2018;26(4):339-50.
168. Jantscher-Krenn E, Aigner J, Reiter B, Kofeler H, Csapo B, Desoye G, et al. Evidence of human milk oligosaccharides in maternal circulation already during pregnancy: a pilot study. Am J Physiol Endocrinol Metab. 2019;316(3):E347-E57.
169. Turroni F, Bottacini F, Foroni E, Mulder I, Kim JH, Zomer A, et al. Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging. Proc Natl Acad Sci U S A. 2010;107(45):19514-9.
170. Marcobal A, Barboza M, Sonnenburg ED, Pudlo N, Martens EC, Desai P, et al. Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. Cell Host Microbe. 2011;10(5):507-14.
171. Katayama T. Host-derived glycans serve as selected nutrients for the gut microbe: human milk oligosaccharides and bifidobacteria. Biosci Biotechnol Biochem. 2016;80(4):621-32.

172. Lugli GA, Duranti S, Milani C, Mancabelli L, Turroni F, Alessandri G, et al. Investigating bifidobacteria and human milk oligosaccharide composition of lactating mothers. FEMS Microbiol Ecol. 2020;96(5).
173. Zabel BE, Gerdes S, Evans KC, Nedveck D, Singles SK, Volk B, et al. Strain-specific strategies of 2'-fucosyllactose, 3-fucosyllactose, and difucosyllactose assimilation by Bifidobacterium longum subsp. infantis Bi-26 and ATCC 15697. Sci Rep. 2020;10(1):15919.
174. Duranti S, Lugli GA, Mancabelli L, Armanini F, Turroni F, James K, et al. Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. Microbiome. 2017;5(1):66.
175. Moya-Gonzalvez EM, Rubio-Del-Campo A, Rodriguez-Diaz J, Yebra MJ. Infant-gut associated Bifidobacterium dentium strains utilize the galactose moiety and release lacto-N-triose from the human milk oligosaccharides lacto-N-tetraose and lacto-N-neotetraose. Sci Rep. 2021;11(1):23328.
176. Nilsson AI, Koskiniemi S, Eriksson S, Kugelberg E, Hinton JC, Andersson DI. Bacterial genome size reduction by experimental evolution. Proc Natl Acad Sci U S A. 2005;102(34):12112-6.
177. Ward RE, Ninonuevo M, Mills DA, Lebrilla CB, German JB. In vitro fermentability of human milk oligosaccharides by several strains of bifidobacteria. Mol Nutr Food Res. 2007;51(11):1398-405.
178. Azad MB, Robertson B, Atakora F, Becker AB, Subbarao P, Moraes TJ, et al. Human Milk Oligosaccharide Concentrations Are Associated with Multiple Fixed and Modifiable Maternal Characteristics, Environmental Factors, and Feeding Practices. The Journal of Nutrition. 2018;148(11):1733-42.

9.0 Appendix

Table S.1 Genomic features of *Bifidobacterium* isolates

SampleID	Genome size	Contigs	GC(%)
E002-T2-1	2384693	39	59.87
E002-T2-2	1918926	42	60.48
E002-T2-3	1918173	51	60.48
E002-T2-4	1918769	42	60.47
E002-T2-5	1918082	44	60.48
E002-T3-1	1918980	71	60.51
E002-T3-2	1919281	46	60.47
E002-T3-3	1919609	47	60.47
E002-T3-4	1919490	46	60.48
E002-T3-5	1918685	45	60.47
E002-T4-1	2325846	29	59.94
E002-T4-2	2313529	28	59.94
E002-T4-3	2352300	34	59.85
E002-T4-4	2316561	30	59.93
E002-T4-5	2315229	28	59.93
E003-T2-2	2212203	63	60
E003-T2-3	2574211	33	58.4
E003-T2-4	1918135	48	60.48
E003-T3-3	1920282	66	60.52
E003-T3-4	1915680	45	60.48
E003-T4-1	2371455	43	59.44
E003-T4-2	2210131	39	59.98
E003-T4-3	2578178	89	58.39
E003-T4-4	2209789	40	59.97
E003-T4-5	2208854	40	59.97
E005-T2-1	2384693	39	59.87
E005-T2-3	2384923	40	59.86
E005-T2-4	2384404	64	59.88
E005-T2-5	2384316	39	59.87
E005-T3-1	2384220	37	59.86
E005-T3-2	2384190	36	59.87
E005-T3-4	2383410	36	59.87
E005-T3-5	2384061	40	59.87
E005-T4-2	1921835	76	60.5
E005-T4-3	1918592	39	60.48
E005-T4-4	1919363	46	60.48
E005-T4-5	1917190	51	60.48
E006-T2-1	2232970	37	60.19

E006-T2-2	2256866	41	60.16
E006-T2-3	2256689	47	60.16
E006-T2-4	2257179	46	60.16
E006-T2-5	2257057	45	60.16
E006-T3-1	2310449	82	60.25
E006-T3-2	2257267	43	60.16
E006-T3-3	1919436	45	60.47
E006-T3-4	2257225	41	60.16
E006-T3-5	2244223	47	60.18
E006-T4-1	2242616	41	60.21
E006-T4-2	2320192	58	60.19
E006-T4-4	2241940	43	60.21
E006-T4-5	1919375	48	60.48
E007-T1G1	1918452	76	60.55
E007-T1G2	1920759	78	60.56
E007-T1G3	1920227	75	60.54
E007-T1G4	1920194	71	60.53
E007-T1G5	1921039	72	60.54
E007-T1R1	1921298	67	60.55
E007-T1R3	1921141	74	60.55
E007-T1R4	1921842	81	60.53
E007-T1R5	1921582	75	60.56
E007-T2G1	2333189	60	59.31
E007-T2G3	2296652	54	59.33
E007-T2R2	2331463	65	59.33
E007-T3G1	2363195	64	59.76
E007-T3G5	2241944	48	59.37
E007-T3R1	2363906	64	59.74
E008-T2-2	2212169	69	60.01
E008-T2-3	2388466	92	60.28
E008-T2-4	1922223	73	60.54
E008-T2-5	2388496	97	60.28
E008-T3-1	2388361	99	60.29
E008-T3-2	2388600	98	60.29
E008-T3-3	2388292	113	60.28
E008-T3-4	2388372	89	60.28
E008-T3-5	2387338	86	60.28
E009-T2-1	1921494	76	60.54
E009-T2-2	1921312	64	60.54
E009-T2-3	1920908	76	60.54
E009-T2-4	1920830	68	60.54
E009-T2-5	1921374	73	60.54
E009-T3-1	1921569	73	60.54

E009-T3-2	1920290	63	60.54
E009-T3-4	3402041	1497	60.08
E009-T3-5	1921472	73	60.55
E009-T4-1	1922083	78	60.54
E009-T4-2	1921234	78	60.54
E009-T4-3	1921246	73	60.54
E009-T4-4	1920021	71	60.54
E009-T4-5	1921158	68	60.53
E010-T1-5	2635467	105	60.21
E010-T2-3	2141115	65	59.67
E010-T2-4	2631080	92	60.21
E010-T3-1	2469374	118	60.37
E010-T3-2	2717991	60	58.36
E010-T4-1	2720902	78	58.38
E013-T2-1	2416729	80	60.18
E013-T2-2	2489581	89	60.22
E013-T2-3	2500676	97	60.22
E013-T2-4	2420376	80	60.2
E013-T2-5	2423223	82	60.11
E013-T3-1	2502464	89	60.22
E013-T3-2	2416361	86	60.19
E013-T3-3	2454893	86	60.13
E013-T3-4	2421240	63	60.17
E013-T3-5	2489664	88	60.22
E013-T4-1	2501913	87	60.22
E013-T4-2	2455264	79	60.14
E013-T4-3	2491322	106	60.24
E013-T4-4	2422954	106	60.18
E013-T4-5	2406914	76	60.14
E014-T3-4	2103714	34	59.34
E014-T3-5	2156224	24	59.45
E014-T4-1	2156208	26	59.44
E014-T4-2	2110212	44	59.33
E014-T4-3	1918878	41	60.48
E014-T4-4	2656934	777	59.54
E015-T2-3	1921773	85	60.55
E015-T2-4	1920561	87	60.53
E015-T2-5	1920834	81	60.55
E015-T3-1	2315275	67	60.13
E015-T3-2	2310492	77	60.14
E015-T3-3	2314790	60	60.14
E015-T3-4	2485301	94	60.21
E015-T3-5	2314844	59	60.13

E030-T1-3	2243056	69	59.27
E030-T1-5	2183942	68	62.88
E030-T2-1	1921897	95	60.52
E030-T2-2	1920164	73	60.53
E030-T2-3	2186399	74	62.89
E030-T2-4	2191572	64	62.89
E030-T2-5	1920462	78	60.54
E030-T3-2	1921856	69	60.53
E030-T3-4	1922278	83	60.54
E030-T3-5	1921074	77	60.54
E030-T4-1	1920801	71	60.54
E030-T4-2	1920156	71	60.53
E030-T4-3	2218484	73	63.01
E030-T4-4	1920672	74	60.54
E030-T4-5	1920702	77	60.53

The following tables can be accessed via this link due to each being large files:
https://1drv.ms/u/s!Atdggw2aUs2SgTHGptdHkiwKcbD7?e=CHBr6P

Table S.2 SNP distance values of *Bifidobacterium longum* isolates

Table S.3 SNP distance values of *Bifidobacterium animalis* isolates

Table S.4 SNP distance values of *Bifidobacterium adolescentis* isolates

Table S.5 SNP distance values of *Bifidobacterium bifidum* isolates

Table S.6 SNP distance values of *Bifidobacterium dentium* isolates